DOCUMENT RESUME

ED 229 359 SP 022 234

AUTHOR	Denton, Cliff; Postlethwaite, Keith
TITLE	The Identification of More Able Pupils in Comprehensive Schools. Final Report. Part II.
INSTITUTION	Oxford Univ. (England). Oxford Educational Research Group.
PUB DATE	Dec 82
NOTE	294p.; For related document, see SP 022 233.
PUB TYPE	Information Analyses (070) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC12 Plus Postage.
DESCRIPTORS	Ability Grouping; *Ability Identification; *Academically Gifted; Checklists; Foreign Countries; *Informal Assessment; Inservice Teacher Education; Secondary Education; Secondary School Teachers; Skill Analysis; Standardized Tests; *Student Characteristics; Student Evaluation; *Teacher Attitudes; *Test Reliability
IDENTIFIERS	England

ABSTRACT

In the second year of a project investigating the ability of secondary school teachers to identify high-ability students, two questions were addressed: what student characteristics influenced teachers' judgments, and why checklists appeared to have little impact on teachers' judgments. A structured approach was developed to study student characteristics to which teachers were sensitive when making judgments on ability. This approach sought to determine the scope of teachers' knowledge of students, and how this knowledge corresponded with judgments based on students' test scores. Checklists were modified to make them more workable. Extensive classroom observations were made to establish if relevant clues to ability could emerge from day-to-day activities. A number of behavioral and motivational influences on teacher judgments, as well as those based on student test scores, were observed. It was concluded that teachers' assessments were sufficiently high for them, combined with refined checklists, to be considered relevant. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED229359

OXFORD EDUCATIONAL RESEARCH GROUP

THE IDENTIFICATION OF MORE ABLE PUPILS IN COMPREHENSIVE SCHOOLS

by

Cliff Denton

and

Keith Postlethwaite

FINAL REPORT

PART II

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. C. J. Denton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented to the
Department of Education and Science
December 1982.

CHAPTER NINE

The Second Year

If we were to leave the research at this point, we could claim to have achieved important insights into the success of the teacher-based identification process as we could expect it to proceed in many classrooms. We could encourage teachers to be more confident in their judgements than they might have been without our work; though we would add that there was evidence to suggest that there were some inaccuracies associated with teacher-based identification in a considerable number of cases. These inaccuracies seemed to be particularly prevalent in some subjects. There was also evidence that overemphasis on some inappropriate characteristics seemed to lead to bias in teachers' judgement. Finally, we have shown that checklists did not have any measurable effect on the identification. Merely asking a teacher to try to recognise appropriate criteria did not work - despite the fact that teachers were quite co-operative in the exercise.

Two major questions were apparent at the end of this work. One concerned the pupil characteristics that influenced teachers judgement. As we have mentioned, we had some interesting insights from the work which was reported in Chapter Four, but these were necessarily limited to information that was related to our test battery. We also had some tantalising clues from our interviews with teachers (Chapter Seven) about the kinds of characteristics that led them to under- and over-estimate potential. However the interviews were deliberately conducted in a very open ended way, and something with a more precise structure was called for if any firm conclusions were to be drawn.

The other question which was unresolved at the end of Year One was that of why the checklists, used in the way in which we had asked teachers to use them, had been of such little value in refining teachers' judgements. One possibility was that the detailed clues to pupils' ability which were referred to by the lists could not be picked up by teachers in the restrictive environment of the classroom because the pupils were not provided with the kinds of work which would enable them to display such clues. Another alternative, however, could be that the way in which the checklists were used was inappropriate.

We would suggest, then, that we cannot provide advice about how improvements in teacher-based identification can be made without further study. It is of considerable importance to distinguish between alternative plausible hypotheses, and to establish whether some of the reasonable hypotheses that we have put forward can in fact be developed into a well-supported structure.

We will see, in the next few chapters, 3 ways in which we have attempted to achieve these aims.

First we developed a much more structured approach to the study of the pupil characteristics to which teachers were sensitive when making their judgements on their pupils' abilities. From this we could discover the scope of teachers' knowledge of pupils and an understanding of which aspects of their knowledge were associated with judgements that were in line with judgements based on the pupils' test scores, and which were distractors, leading teachers to over or underestimate pupils in comparison with their test scores.

The details of this method are discussed in Chapter Ten and the results are to be found in Chapter Eleven.

Secondly we gave checklists a further chance. By modifying their methods of use through negotiation with the teachers, we sought to make them more workable.

Thirdly we made extensive classroom observations to see if we could establish if the relevant clues to ability could emerge from the day to day activity of pupils in class. It may have been that because of the many limitations of time and opportunity it was not possible for teachers to have the detailed knowledge of children that is required if checklists are to be of help. This intensive activity was possible in just one subject.

The more detailed study of checklists is described in Chapter Twelve.

Thus we have attempted to come away from the computer output into the classroom so that we could add the necessary evidence which would allow the reader to make a more balanced judgement about the problems and possibilities associated with teacher-based identification strategies.

4

CHAPTER TEN

An Approach to the Understanding of Teachers' Views of Pupils based on Personal Construct Theory

In Chapter Seven we described a series of interviews with teachers through which we gained some interesting insights into the kinds of pupil characteristics that they appeared to notice. We also reported that we were able to form a view that some of these characteristics seemed to be associated with teachers' overestimation of pupils' potential, while others seemed to be associated with underestimation of pupils. In that earlier chapter we stressed the tentative nature of these insights. We drew attention to the fact that we could not be certain that we had assembled a complete picture of the range of characteristics that were noticed by teachers. We also pointed out that we had no basis on which to estimate whether the comments made by a teacher in speaking about one individual pupil were part of that teacher's way of looking at all pupils, or whether the comment was very specific to the individual pupil who was being discussed. To these, and the other limitations of the interview technique, we would add one further comment: the procedure generated interesting hypotheses concerning the factors which were associated with under- and over-estimation of potential, but it provided no grounds on which the validity of these hypotheses could be tested. This chapter, and the next, describe the background, and then the results of a more detailed exploration of these issues.

The theoretical basis that we have adopted for this more detailed exploration of the way in which teachers perceive their pupils is that of Personal Construct Theory (Kelly, 1955). Pope and Keen (1981) draw attention to the relevance of this theory for the study of many kinds of problem in education and point out that little effort has as yet been expended on the particular topic of teacher assessment of students despite the fact that, in their view, the theory had much to offer in this area.

In this context of teacher assessment, what the theory suggests is that the teacher makes sense of his interactions with a wide range of pupils by setting up a number of reference axes (or constructs) in his mind and placing each pupil at some point in the multidimensional space defined by these axes. For example: a given teacher may think in terms of a three dimensional space consisting of the axes 'quiet-noisy', 'neat-untidy' and

5

'able-not able'. He may then place an individual pupil at the left hand extreme of the axis labelled 'quiet-noisy', at the centre of the axis labelled 'neat-untidy' and near the right hand end of the 'able-not able' axis. This placement of the pupil in the teacher's three dimensional map enables the teacher to make sense of his previous experience of this pupil and to predict how the pupil may behave in new situations (e.g. since the pupil is 'coded' as quiet and not very able, he will not be expected to make many insightful contributions in a new situation such as a cut and thrust debate about a complicated issue).

The axes set up by an individual teacher may or may not be orthogonal to each other. If they are orthogonal, the placement of a pupil on one axis can be entirely independant of his placement on the other axes. If they are not orthogonal, the teacher is using a map in which one characteristic is taken as being correlated with another (If our teacher, in the example quoted above, regarded most quiet people as able, the 'quiet-noisy' and the 'able-not able' axes would not be orthogonal in that teacher's map.)

Using such a model, the problem of understanding how teachers go about the business of identifying pupils as more able can be thought of as consisting of several separate steps: first we need to know the nature of the individual axes - or constructs - that the teacher uses; secondly we need to know the structure of the relationship between these individual constructs (e.g. which constructs are orthogonal); thirdly we need to know where the teacher places each pupil in the map defined by these constructs; and finally we need to know how the teacher interprets the pupil's position in this map in trying to come to an overall decision about whether or not that particular pupil should be classified as more able. We will consider some of these stages later, but we will first continue with some further general comments about the decision to use Personal Construct Theory as the basis of this part of the work.

We suggest that the theory provides a particularly useful context for our detailed study of the way in which teachers come to a decision that a particular group of pupils have high ability in a given subject. This can be seen, first, from the fact that the theory is essentially a Personal Construct Theory: i.e. it demands that we look at the nature of the constructs used by each individual teacher, and it therefore allows us to continue to deal with the variety that we have already found to be present in the way in which different teachers look at their pupils. Secondly, by encouraging us to look at the place each pupil occupies in the teacher's construct space, it provides a framework for the collection of quantitative data which, nevertheless, is still related to the teacher's own personal system of mapping. This quantitative data is more amenable to analysis than, for example, the qualitative descriptions of pupils elicited in the interviews. Thirdly, the theory draws our attention to the fact there there may be relationships between the teacher's judgement of pupils on one construct and his judgement of those pupils on some other constructs. This is consistent with our wish to explore more fully the nature of the association between particular characteristics noticed by the teacher, and that teacher's overall assessment of the potential of a pupil.

Having established the relevance of the personal construct viewpoint to our present concerns it is useful to consider where the adoption of such a theory places us in relation to the range of views which are held about the nature of person perception. It must immediately be admitted that it places us firmly in the school of thought which regards person perception as a matter of manipulation of component pieces of information which ultimately leads the perceiver, rather in the manner of Sherlock Holmes, to some overall judgement about the perceived. In contrast, there is a second broad school that regards person perception as essentially an intuitive process. This position is described by Cook (1979) as one in which people understand others 'by an act of intuition encompassing their personality as a whole'.

This school of thought, which puts forward similar views to those advanced by Gestalt psychologists, suggests therefore that we respond to the 'meaningful whole' of others and that this 'whole' can never be adequately represented as a sum of component parts. This is the theoretical framework into which ideas such as 'feminine intuition' can be fitted, and in some circumstances it certainly seems to ring true. It accounts, for example, for the situation in which an interviewer 'just knows' that the candidate is going to be right for the job.

Despite the superficial attractiveness of the intuition-based view of person perception, it has a major shortcoming compared with the inference-based model which we have chosen to adopt. This shortcoming arises out of the fact that it is extremely difficult to see what kind of detailed investigation could (or should) be undertaken if all that we can say about teachers' perceptions of pupils is that they 'just knew'. It is therefore a model which leaves one with no indication of which paths to follow in trying to help teachers to be able to identify able pupils more accurately. This is the basis of our decision to discard this intuition-based model. In doing so we can draw support from the views expressed by Cattell (1937) who described the intuition-based model as obscure and unscientific and argued that it:

> contributes nothing except a magic word the purpose of which is to cry 'Halt' to the irreverence of further investigation.

We therefore suggest that there is considerable justification for our recourse to the inference-based model. However it would be wise to draw attention immediately to the debate which exists about the precise nature of the process through which the inferences are made. In doing so we hope to provide some theoretical basis for the kinds of analysis which we have undertaken and to draw attention to one further advantage of the Personal Construct Theory approach (i.e. that it does not commit us to one particular view of the process of inference but that it provides, instead, data which can be analysed in a variety of ways, drawing upon different theoretical models as appropriate).

One much discussed model of the process of inference is that offered by Sarbin (1944) who makes an analogy with the logic of the classical syllogism:

All X are Y.
Z is X.
So Z is Y.

Sarbin suggests that this model has to be modified to the extent that the major, and sometimes also the minor premises are, in the person perception context, expressed in probabilistic terms. For example:

80% of sports car drivers are aggressive.
John is a sports car driver.
So there is a 0.80 chance that John is aggressive.

Considerable debate has been generated by this model (Meehl,1954; Sarbin et al, 1960; Meehl, 1961). In summarising this debate, Cook (1979) makes the point that 'nothing follows logically from the major premise "Some X are Y"'. This view would seem to be consistent with that of Allport, which is quoted - with, it seems, some sympathy -by Meehl (1954). Their argument rests on the idea that it is a nonsequitur to conclude that 'there is a 0.80 chance that John is aggressive', for, they say, we know that the real position is that John is either aggressive (in which case there is a chance of 1.00 that he is aggressive), or he is not (in which case the chance is 0). These are the only possible positions and in both of them the figure of 0.80 has no real meaning. They suggest that the information provided is not sufficient to enable us to pass any judgement on John's aggressiveness on the basis of reasoning of this kind. The implication in this line of argument is that, since we do come to conclusions about people on the basis of such data, we must actually structure our reasoning in some other way. We will present alternative models (including that due to Meehl) shortly, but we feel that it is worth looking a little more closely at the substance of this argument before moving on.

While we accept that, as far as it goes, the point of view expressed in the previous paragraph is sound, we would suggest that it leads one to undervalue Sarbin's model. Surely the point is that, in the absence of any of the further information necessary for the certain classification of John, the most sensible thing that we can say about him is that he (like 80% of sports car drivers) will be aggressive, and that the figure of 80% gives us a basis for assessing the odds that we will be right in making such a judgement. We would therefore suggest that the reasoning may run as follows:

80% of sports car drivers are aggressive.
John is a sports car driver.
So in the absence of further data I will assume that John is aggressive
And the odds are 4:1 that I shall be proved right.

This model has moved some way from the classical syllogism but it is still a model based on the categorisation of the individual, and on inferences drawn from what is known about the category into which the individual is put.

We turn now to Meehl's own alternative model of the process of inference which is essentially a correlation model. Meehl (1961) proposes that the perceiver does not think of placing the perceived person into categories and of indulging in something similar to syllogistic reasoning on the basis of that categorisation. Rather, he suggests, the perceiver reasons on the basis of the degree to which the perceived person has some property 'x', and on the basis of the strength of the correlation between this property 'x' and some other property 'y', so as to come to some view on the degree to which the person can be expected to have this second property 'y'.

Because, as we have suggested above, there is no watertight logical reason to discard Sarbin's model, we suggest that the choice between the Sarbin and Meehl models is by no means straightforward when one is considering the simple situation in which a person is reasoning on the basis of a single cue. However Meehl makes a strong case that his model, when generalised from simple correlation to multiple regression, is far more successful than the syllogistic model in describing the situation in which a perceiver is trying to handle a large number of partially correlated cues. Indeed it is clear that any attempt to express such a problem as a series of syllogisms would be enormously complex.

It would seem that the data which we have collected within the personal construct framework (consisting as it does of ratings of pupils on each of their teacher's constructs) is very amenable to analysis in terms of this generalised version of Meehl's correlational model for the concept of 'the degree to which the pupil has some property 'x'' is clearly represented in such data. However, as we hinted earlier, one advantage of the personal construct approach is that it does not limit us to this model of the process of inference. Cook (1979) describes a number of alternative models for the way in which people handle multiple cues. He suggests that people may add, or average information on different aspects of another person to come to an overall judgement about that person, and that inconsistent information may be discounted. He suggests that in the process of discounting some parts of the set of information, positive information tends to be overlooked more readily than negative information.

He also suggests that people may make use of favourite cues and of disqualifying cues. In our context, this last suggestion would imply that the teacher may be sensitive to a particular pupil characteristic that he/she regards as so firmly linked to high ability that a pupil displaying such a characteristic will almost always be described as more able whatever any further information on the pupil might indicate (such a characteristic would be operating as a favourite cue for that teacher).

Similarly the teacher may be sensitive to another characteristic which is so firmly linked in the teacher's mind with low ability that pupils displaying this characteristic will almost always be denied membership of the more able group whatever other information is available (such a characteristic would be operating as a disqualifying cue).

The concepts of favourite and disqualifying cues (which we would see, incidentally as more consistent with the syllogistic, rather than the correlational model of the process of inference) are of particular interest to us as we already have some indication from the general interviews with teachers that they seemed to use some characteristics in this way. If this could be confirmed by more detailed study valuable insights could be gained into why teachers overlook some pupils whom the tests suggest are of high ability, and why they choose others whom the tests suggest are not of high ability. Once the nature of these problems is understood, steps could clearly be taken to help teachers to be more accurate in their overall judgements.

We have therefore given some thought to the ways in which personal construct data could be used to investigate these ideas. We have considered three major topics: first how the concepts of favourite and disqualifying cues could be assimilated into the personal construst framework; secondly how one might try to elicit the nature of the cues which a teacher might be using as favourite, or disqualifying cues; and thirdly how one might test the hypothesis that these cues were actually being used in these ways. In the rest of this chapter we discuss the first two points in some detail. The third will be discussed in the next chapter, where the results of this work are presented.

The relationship between the concepts of favourite and disqualifying cues, and constructs is essentially a simple one, but we will discuss it here as it allows us to introduce some terminology that will be helpful elsewhere in the argument. We begin by discussing the disqualifying cue.

We suggest that the idea of a disqualifying cue may be helpful in trying to understand why teachers overlook some pupils who, according to the tests, have high potential for the subject. We could call such pupils 'misses' and could label them as ($T\bar{P}$) (i.e. nominated by the test (T) but not by the teacher - i.e. not of high perceived potential ($\bar{P}$)). Because we wished to continue to take account of test error in identifying pupils who fitted this specification we decided that we would not count any pupil who fell within the top 10% group as a (T) pupil (a pupil nominated by the test), but would restrict this label to pupils who were members of the A Group (see Chapter Three). The idea that a teacher is influenced by a disqualifying cue in forming an overall judgement about such pupils suggests that he/she may have a single construct (or group of constructs) which have one pole which is so firmly a negative correlate of that teacher's concept of high potential that pupils placed at this pole will tend to be overlooked (labelled as ($\bar{P}$)) irrespective of their position in the rest of the teacher's construct space.

Similarly, the idea of a favourite cue may be helpful in trying to understand why teachers choose some pupils who, according to the test, do not have high potential for the subject. We could call such pupils 'false hits' and could label them as ($\bar{T}P$) (i.e. not nominated by the test ($\bar{T}$) but nominated by the teacher - i.e. of high perceived potential (P)). Again, to take account of test error we decided to restrict the label ($\bar{T}$) to those pupils who fell into the C Group. The idea that a teacher may be influenced by a favourite cue in forming judgements about these pupils suggests that he/she may make use of a single construct (or group of constructs) which have one pole that is so closely associated with that teacher's view of high potential in the subject that pupils placed at this pole will tend to be chosen (labelled as (P)) irrespective of their position in the rest of the teacher's construct space.

There is no guarentee that disqualifying and favourite cues will lie at opposite poles of the same construct (e.g. a teacher's disqualifying cue may be 'uncouth behaviour' but his/her favourite cue may be 'good at arithmetic') so care must be taken to ensure that both kinds of construct can emerge from a personal construct interview.

Before turning to our second problem of how the constructs that involved favourite and disqualifying cues could be elicited, we will give a very brief description of the usual methodology employed in personal construct studies.

The classical approach to the problem of how to find out what kinds of constructs an individual uses in structuring his or her experience of a particular group of elements (e.g. people in the individual's family, people he/she works with, types of music, styles of writing or, as in our case, pupils taught by the individual) is the approach known as triadic elicitation. In this procedure a list of relevant elements is produced and groups of three elements from this list are constructed. Each group of three elements (each triad) is then presented to the individual who is asked to say in what way two of the elements are alike, and in what way the third is different from the other two. For example, in a study of reactions to types of music we might present a triad consisting of Offenbach, Gilbert and Sullivan and Wagner operas. From such a triad we might elicit the comment that Offenbach and Gilbert and Sullivan operas are light, whereas Wagner's operas differ from the others by being serious. We would then know that one construct (or reference axis) used, in relation to music, by the person to whom we had offered the triad, was the construct: 'serious -light'. We would continue by offering other similar triads, so that we could begin to build up a full picture of how our subject thinks about types of music. This procedure, together with other aspects of the use of Personal Construct methodology, is described briefly by Hall (1978), and is discussed in detail by Fransella and Bannister (1977).

In preparation for the discussion of how this procedure must be modified to suit our main purpose it is useful to note that personal construct studies are often used to provide descriptions of the way in which people view some aspect of their experience, or as an aid to diagnosis, for example in the treatment of psychiatric patients. In such circumstances, the overriding consideration governing the elicitation procedure is that it should provide a representative map (or ideally a complete map) of the whole range of constructs used by the individual in relation to a particular kind of element. This consideration implies that the triads we present to the individual should be representative of all possible triads that could be constructed from the relevant list of elements; and this, in turn implies that the most suitable way of choosing the three elements for each triad is on the basis of random selection. This advice is clearly stated by Yorke (1978) who advises that:

> without any strong justification for choosing to use particular triads it is probably preferable to assign elements randomly.

Certainly one purpose of our personal construct work was to provide a description of the kinds of pupil characteristics that were noticed by teachers so that we could extend the general insights gained from the interview study; but this was largely incidental to the second, and more important purpose, which was, as we have explained above, to explore the idea that some constructs might be associated with favourite cues, and others with disqualifying cues.

In connection with the first purpose a representative mapping was required and, therefore, the presentation of randomly chosen triads was called for. However the same considerations do not apply in connection with the second purpose. The over-riding consideration there, was that of encouraging the teacher to mention the constructs that were associated with any favourite and disqualifying cues that he/she might use. Clearly the appropriate constructs would only emerge if a $(T\bar{P})$ pupil (one who was overlooked despite high test score) or a $(\bar{T}P)$ pupil (one who was chosen despite low test score) was presented, in a triad, with appropriate contrasting pupils.

By reference to Chapter Three the reader will be able to see that because of the considerable degree of match between test and teacher judgement of pupils, there were relatively few pupils who could be labelled as $(T\bar{P})$ or $(\bar{T}P)$ in any one subject in any one school. Therefore, only if large numbers of triads were presented to each teacher, could one rely upon the random selection of pupils for each triad to generate the appropriate groupings. This recourse to large numbers of triads was itself rather unattractive, partly because it would have been unacceptably demanding on the time of the teachers, but more seriously because the presentation of large numbers of triads would have been exhausting for both parties involved in the elicitation interview (Pope and Keen, 1981). As a result, the long interviews might be expected to become more and more unrewarding as they continued. We were therefore drawn back to a position not unlike that adopted by Kelly himself, in his seminal work on Personal Construct Theory (Kelly, 1955, p.273 et seq). He suggested using, and gave examples of, carefully chosen triads which could be offered to elicit constructs of a particular kind. Of course, as the kinds of constructs in which we were interested differed from those of which Kelly was writing, the details of the formulae for the construction of these special, or 'critical' triads were necessarily different in our case. Nevertheless the principle is the same. This idea has also been suggested by Easterby-Smith (1981) though he provides nothing more than the hint that it might be useful to present triads which will 'bring out the greatest contrast in the elements available'.

As an example of the kind of critical triad that was relevant to our work, consider the following group of pupils (T$\bar{P}$; T$\bar{P}$; TP). These three pupils had similar high scores on the test measure of potential and so were all classified as (T). One also had high perceived potential (P), but two were not regarded as members of the top 10% group by their teacher ($\bar{P}$). If, in the elicitation interview, the teacher grouped the pupils thus (T$\bar{P}$; T$\bar{P}$) and (TP), one might expect that the resulting construct would be an important part of the way in which that teacher was led towards making errors of under-estimation, and that it contributed to the 'misses' that he/she made. One might also suspect that one pole of this construct (that describing the pair of (T$\bar{P}$) pupils) might have been operating as a disqualifying cue for that teacher. Clearly other distinctions between the pupils in the triad might override this distinction based on test and teacher perceptions of potential and one would not, therefore, expect all triads of this type to be split by the teacher in the way which was suggested above. Nevertheless when such a split was made, it was possible to make the kind of classification of the construct that we have indicated above.

A very important point emerges from this paragraph. Not only does the presentation of carefully designed 'critical' triads encourage the teacher to mention the particular kinds of construct in which we were especially interested, it also allows us to make an a priori classification of the resulting construct. The appropriate constructs might emerge if enough randomly designed triads were presented but no a priori classification would then be possible. Because the 'critical' triad approach gives rise to this possibility of forming classifications of triads that can later be tested to see if it holds up, it effectively shifts the technique from one of description and diagnosis to one of experimental exploration of the topic. We believe that this represents a fairly novel way of putting personal construct methodology to work.

Having established the idea of 'critical triads' we turn now to the question of how many different kinds of critical triad can be designed with the kind of data which we have in our study. For this part of the work we can regard our problem as one involving two-way classification (pupils are labelled as more able or not by the teacher, and as more able or not by the test). There are therefore four different types of pupil (TP), ($\bar{T}$P), (T$\bar{P}$) and ($\bar{T}\bar{P}$). We can clearly make up four types of triad in which all three pupils are the same, and four types in which all three pupils are different from each other (e.g. TP, $\bar{T}P$, $T\bar{P}$). In each of these cases a priori classification of the constructs which emerged from the triads would be impossible. However constructs emerging from the other twelve triads (in which one pupil differed from the other two) could be classified on an a priori basis. These triads are listed in Table 10.1, together with the appropriate classification.

TABLE 10.1

POSSIBLE CRITICAL TRIADS

Triad type A1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, TP

Triad type A2: TP, TP, $\bar{T}\bar{P}$

These triads give rise to constructs on which the teacher's judgement categorised the pupils in the same way as the test.

Triad type B1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, $\bar{T}$P

Triad type B2: $\bar{T}$P, $\bar{T}$P, $\bar{T}\bar{P}$

These triads give rise to constructs on which the teacher separated pupils with similar low test score. The pole at which $\bar{T}$P was placed may have been a favourite cue.

Triad type C1: TP, TP, T$\bar{P}$

Triad type C2: T$\bar{P}$, T$\bar{P}$, TP

These triads give rise to constructs on which the teacher separated pupils of similar high test score. The pole at which TP was placed may have been a disqualifying cue.

Triad type D1: T$\bar{P}$, T$\bar{P}$, $\bar{T}$P

Triad type D2: $\bar{T}$P, $\bar{T}$P, T$\bar{P}$

These triads give rise to constructs on which the teachers' perceptions lead to reverse assessments from those based on the test. Favourite and disqualifying cues may be found at the opposite poles of these constructs.

Triad type E1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, T$\bar{P}$

Triad type E2: T$\bar{P}$, T$\bar{P}$, $\bar{T}\bar{P}$

In these triads $\bar{T}$ and T are distinguished by the teacher but the characteristics displayed by the TP pupils are undervalued as indicators of high potential. The TP pole may again be a disqualifying cue.

Triad type F1: TP, TP, $\bar{T}$P

Triad type F2: $\bar{T}$P, $\bar{T}$P, TP

In these triads T and $\bar{T}$ are distinguished by the teacher but the characteristic displayed by the $\bar{T}$P pupils is undervalued by the teacher as an indicator of low potential. The TP pole may again be a favourite cue.

Throughout the table the construct classification only applies if the triad is split so that the third member is identified as the 'odd one out'.

It must be recognised that an elicitation procedure based only on these critical triads may produce a distorted grid, limited by our own perception of the nature of the problem (Yorke, 1978). Our major response to this criticism is that we have traded a representative survey of teachers' constructs for an experimental study of constructs which behave as favourite and as disqualifying cues. We also note that, through critical triads A1 and A2 we get insights into the nature of the constructs which lead teachers to make classifications of pupils which are in line with those made by the tests. However, to go some way to overcoming this criticism in a more direct manner, some triads made up of randomly chosen pupils were included in our elicitation interview together with triads fitting the specifications listed in Table 10.1. In presenting the triads, critical and randomly chosen triads were intermixed, and the overall order of presentation was chosen at random.

Names of pupils to form the critical triads were selected in the following way. The pupils taught by a given teacher were each labelled as TP, T̄P, etc. When a pupil of a particular kind (e.g. TP) was needed in a triad, a name was chosen at random from the set of pupils of that kind who were taught by that teacher. When triads suiting all of the specifications in Table 10.1 had been constructed, they were listed in the order of presentation. They were then inspected and some changes were made, if necessary, to avoid repetition of pairs of names in successive triads. (Bender (1974) found that there was a significantly greater tendency to produce important constructs when successive triads were varied by two elements at a time, rather than one at a time.)

The intention was that the elicitation interview should present each teacher with fifteen different triads. Some teachers, who felt that they had exhausted their repertoire of constructs before this point had been reached, actually handled fewer than fifteen triads. Others dealt with several more. The time available for the interviews (normally something of the order of 90 minutes) enabled this scale of activity to be conducted in a busy, but not unduly rushed fashion. It is interesting to note that all of the teachers who went beyond fifteen triads felt that they had nothing extra to offer after dealing with at most five additional triads. It seems that it would indeed have been of little value to attempt to work with very large numbers of randomly chosen triads.

In the next chapter we present the results obtained from this study.

The work was done in detail in physics and on a smaller scale in the other subjects. We therefore present the physics results first to give a clear picture of the various stages of the analysis and interpretation of the findings. We then summarise the outcome of the work in the other subject areas.

CHAPTER ELEVEN

Results of the Personal Construct Study

In this chapter we give an account of the findings which have arisen to date, out of the personal construct work that was described in the previous chapter. It must be emphasised that this has proved to be a very fertile line of enquiry and that the resulting data lends itself to a much more extensive analysis than we have been able to undertake within the constraints of the present project.

The main aim of the work which has so far been completed has been to gain insights into the kinds of constructs which teachers use in assessing pupils, and to indicate:

i) which of these constructs seemed to be associated with accurate overall judgements of pupils ability;

ii) which seemed to have poles that operated as favourite cues, and were therefore associated with over-estimation of ability;

iii) which seemed to have poles which operated as disqualifying cues, and were therefore associated with an under-estimation of ability.

Some data relating to these main aims is presented for all four subjects. However, the demands on time made by other parts of the project led us to undertake the majority of the fieldwork in physics. We have therefore taken the analysis of the physics data rather further than the analysis of the data in the other subjects. We will therefore begin by reporting this work in physics. It is perhaps worth emphasising that even in this subject, there is much more that could be done (and that we hope will be done) in the future.

11.1 Basic results in Physics

In physics, 182 triads were presented in all to 11 teachers in 7 schools: 140 of these triads were critical triads; 42 were randomly chosen ones. In the case of 19 triads (12 critical triads, and 7 randomly chosen ones) the teacher found it impossible to divide the pupils into a group of two, and a contrasting odd-man-out. It is interesting to note that 8% of the critical triads proved to be impossible for the teacher to split, whereas 17% of the random triads suffered this fate. This is not surprising, in that the critical triads were specifically chosen to contain pupils who differed in some way, but it does suggest that the use of critical triads did, as predicted, add to the efficiency of the elicitation procedure[1].

Of the 128 critical triads that were split by the teachers, 80 (62.5%) were split as one might have expected from the design of the triad. It is difficult to draw any firm conclusions from this figure, as not all of the triad divisions made by teachers can be thought of as independant of one another. (For example, the same pupils sometimes appeared in several of the triads which were offered to a given teacher.) However, the number of triads that were divided in the expected way, was nearly twice the number that one would expect to be divided in that way if the teachers' divisions of the triads had been completely uninfluenced by the design of those triads. In most cases, this high number of 'expected' splits was confirmed when each teacher's choices were considered separately. The results are summarised in Table 11.1.

TABLE 11.1

CRITICAL TRIAD SPLITS ANALYSED BY TEACHER[2]

TEACHER	NUMBER SPLIT	NUMBER SPLIT AS EXPECTED
03	10	9
04M	10	7
04H	12	10
05	11	6
07G	15	8
07B	12	7
08	9	6
09	14	9
10	9	5
11	13	4
12	13	9
Overall	128	80

For comparison, the split of random triads was also analysed. There was, as one would expect, no evidence that the way in which these triads had been split had been biased towards the choice of any particular pupil in the triad (e.g. the first pupil).

Finally, each type of triad identified in Table 10.1 was analysed separately. The results are shown in Table 11.2.

TABLE 11.2

CRITICAL TRIAD SPLITS ANALYSED BY TRIAD TYPE

TYPE	NUMBER SPLIT	NUMBER SPLIT AS EXPECTED	SIGNIFICANCE (p)
A1	11	9	0.001
A2	10	6	0.12
B1	7	5	0.04
B2	12	8	0.02
C1	11	9	0.001
C2	12	6	0.18
D1	10	7	0.02
D2	13	7	0.10
E1	12	8	0.02
E2	7	4	0.17
F1	18	8	0.20[3]
F2	4	3	0.11

In general each teacher handled one triad of each type, so the splits made in each case within a particular type of triad, could be regarded as independent of the other cases within that type of triad. One can therefore calculate the probability of achieving exactly the observed number of 'expected splits' within each triad type, on the basis of chance alone. A simple summing of such probabilities then enables one to calculate the probability of making this number, or more expected splits on the basis of chance alone[4]. This figure can be interpreted as a significance level, indicating the significance of the deviation of the observed result from that which would have been obtained purely on the basis of random selection from the triads.

In six cases the results in this table were significant at the 0.05 (5%) level. In three further cases the probability that the observed number of 'expected splits' would arise by chance was approximately 0.1, and in no case did it rise above 0.2. Of course, when small numbers are associated with each significance test, the results (both those that indicate significance and those that do not) should be treated with caution. A change of just 1 in the number of splits made as expected, can make a great difference to the significance level[5]. However, the results do provide some evidence that teachers did seem to be influenced by the design of the triads. More often than could be expected if they were splitting the triads at random, they chose, as the odd-man-out, the pupil whose specification (in terms of T, $\bar{T}$, P and $\bar{P}$) would lead us to expect him to be chosen.

Table 10.1 provided an a priori classification of the constructs which could be elicited from particular triads if they were split in the 'expected' way. In the following tables we list such constructs under the relevant a priori classifications. We also include, in the Table 11.9, a list of constructs arising from random triads, and from triads that were not split in the 'expected' way. A short discussion of the results follows each table.

TABLE 11.3

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT TP pole		$\overline{TP}$ pole
03	Good overall physics ability	-	Poor overall physics ability
04H	Asks perceptive questions	-	Does not ask perceptive questions
04M	Prepared to answer questions in class	-	Tends not to answer questions in class
05 (a)	Mature formal written work	-	Casual written expression.
05 (b)	Doesn't exert self	-	Makes most of opportunities
07B	Able in mathematics	-	Not able in mathematics
07G	High ability	-	Low ability
08 (a)	Shows flair for subject	-	No flair shown
08 (b)	High performance	-	Low performance
09	Puts in a great deal of effort	-	Jogs along with adequate effort
10	High mathematics ability	-	Low mathematics ability
11	-		-
12 (a)	At home with subject	-	Finds physics difficult
12 (b)	Has natural ability	-	Has to try

These constructs indicate some of the dimensions on which the teachers assessed their pupils. The right hand poles of these constructs indicated characteristics which teachers noticed in pupils who were not judged, by them, to be more able. These were pupils who were also classified by the tests as not more able. In contrast, teachers offered the left hand poles of these constructs as descriptions of pupils who were assessed by the tests, and by the teacher, as more able. The left and right hand poles of these constructs could therefore be regarded as pupil characteristics to which the teachers were sensitive, and which they themselves associated with, or interpreted as signs of, high and low potential respectively. The pupils' test results confirmed the teachers' overall judgement of these pupils, and provided support for the view that the teachers' interpretation of the implications of the various characteristics as indicators of potential, was appropriate.

Present performance in the subject or in a specific part of it (i.e. mathematics), or high general ability (e.g. as exemplified by mature written work) seem to have been the main characteristics that were appropriately interpreted as signs of high potential. 'Flair' and 'natural ability' also seem to have been recognised. Construct (b) from Teacher 05 calls for some comment as at first sight it seems a little odd. It is interesting to speculate that the pupil 'Doesn't exert self' because he 'Doesn't need to exert self', in which case this seemingly reversed construct would fall in line with the two constructs from Teacher 12, and perhaps, construct (a) from Teacher 08.

In Table 11.4 we turn our attention to the constructs elicited when teachers split triads of Type B1 and B2 in the expected way.

TABLE 11.4

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	**CONSTRUCT**		
	$\bar{T}P$ pole		$\bar{T}\bar{P}$ pole
03 (a)	Absorbs day to day work	-	Doesn't absorb day to day work
03 (b)	Enthusiastic	-	Not enthusiastic
04H	Confident in practical work	-	Hesitant in practical work
04M(a)	Concentrates	-	Unable to concentrate
04M(b)	High overall ability in physics	-	Low overall ability in physics
05	Good, conforming behaviour	-	Non-conformist behaviour
07B	Good test results	-	Low test results
07C	-	-	-
08	High performance	-	Low performance
09	Applies him/herself in lessons	-	No drive in lessons
10	Conscientious	-	Very lazy
11	Good mathematically	-	Not good mathematically
12	Doesn't ask questions	-	Asks questions

On these constructs, teachers were separating pupils of similar, low test score. The pupils who were described in the terms set out under the $\bar{T}P$ pole, were pupils who were regarded as more able by their teachers despite the fact that their test scores were low[6]. An interesting feature of the descriptions under this pole - descriptions which we argued earlier might be acting like 'favourite cues' for a teacher - is that there is a clear reference to 'model pupil behaviour' (03(a), 03(b), 04H, 04M(a), 05, 09 10). Behaviour of this kind may of course be associated with high potential in some cases, but the indication from this work is that teachers may have interpreted it in this way even when the test score would suggest that this interpretation was unduly optimistic.

It is interesting to speculate on what was happening where teachers labelled a pupil with low test score as a pupil of 'high ability' (04M(b), 08). It may be that the process of interpretation of cues was happening so automatically at a non-verbal level, that the teacher was not aware that what he/she labelled as 'high ability' was actually observed as, for example 'application' or 'enthusiasm'. An alternative explanation is, of course, that pupils so labelled by test and teacher were pupils whose test scores were in error. Although attempts were made, by choosing pupils with extreme test scores, to reduce the impact of this test error on the triadic elicitation procedure, one must recognise that some individual pupils would have had a true potential which was considerably different from that indicated by the test score.

Table 11.5 indicates the constructs elicited from triads of Types C1 and C2.

TABLE 11.5

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT $\overline{TP}$ pole	 TP pole
03 (a)	Scatterbrain in work	Well organised in work
03 (b)	Low accuracy in written work in physics	High accuracy in written in physics
04H(a)	Low quality content in written work	High quality content in written work
04H(b)	Slow to grasp ideas	Quick to grasp ideas
04H	Low overall ability in physics	High overall ability in physics
05	Not assertive	Makes presence felt (in class)
07B	Less able in written work	Able in written work
07G	Low performance	High performance
08	-	-
09	Questions have to be directed to him to involve him	Contributes without specifically being asked
10 (a)	Quiet worker	Interested in and enthusiastic for the subject
10 (b)	Reluctant to hand in homework	Homework usually in on time
11	Shows poor understanding of physics	Shows good understanding of physics
12 (a)	Finds concepts (in physics) difficult	Good grasp of concepts
12 (b)	Careless work	Careful work

On these constructs teachers separated pupils of similar, high test score. The pupils described under the $\overline{TP}$ pole were pupils who were not regarded as more able by their teachers, despite their high test scores[6].

Some interesting possible disqualifying cues can be seen at the left hand pole of the constructs listed in Table 11.5. The picture that emerges is that teachers interpreted qualities such as disorganisation (03a, 10b), carelessness (03b, 12b,), withdrawn behaviour (05, 09, 10a) and slowness

(04H, b) as indicative of low potential. Where constructs such as 'low-high performance' were used, the discussion set out in the previous section would seem to be relevant again.

Table 11.6 indicates constructs elicited from triads of types D1 and D2.

TABLE 11.6

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT	
	$\bar{T}P$ pole	$T\bar{P}$ pole
03	High self motivation	- Low self motivation
04H(a)	Careful presentation of written work	- Poor presentation written work
04H(b)	Quick to grasp ideas	- Slow to grasp ideas
04M(a)	Low overall ability	- High overall ability
04M(b)	Tends not to answer questions	- Prepared to answer questions
05	-	-
07B	-	-
07G	Low ability	- High ability
08	Committed to the subject	- Not committed to the subject
09 (a)	Made good progress	- Fails to show early promise
09 (b)	High general standards e.g. of politeness	- Low general standards
10	-	-
11	Work improving	- Started well but not maintained progress
12 (a)	Serious about work	- Casual attitude to work
12 (b)	Passive in lessons	- Active in lessons

The contrasts required of the teacher in handling these triads were different from those required in relation to Tables 11.4 and 11.5, so different constructs could be expected. However, the descriptions at the left hand poles can be thought of as contributing further to the list of perceived characteristics which were associated with over-estimation of potential, and those at the right hand pole to the list of perceived characteristics associated with under-estimation of potential. It is, therefore, interesting to see echoes in this table of the findings set out in Tables 11.4 and 11.5.

In two of the constructs in Table 11.6 (04Ma and 07G), the teachers' judgements when dividing the triad seem to be inconsistent with the teachers' judgement when assessing overall potential (P). One can only assume that the teachers' views of the relevant pupils changed between making up the lists on which the labels P or $\bar{P}$ were based, and taking part in the elicitation interview.

It is interesting that two teachers (09, 11) described pupils with high test score but low perceived potential as 'failing to show early promise'. Since the test assessment of their potential, and the teachers implied early assessment of high promise, were in agreement, it is interesting to ask why such pupils fail to fulfil this promise and cease to be labelled as more able by their teachers. It may be that these able pupils were quick to respond to new challenges (such as a new course of study at school or a novel objective test) but also quick to disengage from these activities when they proved to be insufficiently demanding. Some of the other behaviours described under this pole (casual work, low general standards, low motivation etc.) might also arise as a consequence of such a disengagement.

In Table 11.7 we set out the constructs elicited from triads of type E1 and E2.

TABLE 11.7

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT	
	$T\bar{P}$ pole	$\bar{T}\bar{P}$ pole
03	Lacks commitment to academic work	Committed to academic work
04H	Poor attitude to school	Good attitude
04M	Arrogant attitude	Accepting of advice and comment
05	Over confident	Doesn't assert self
07B	-	-
07G	High ability	Low ability
08	-	-
09 (a)	Jogs along - adequate effort	Doesn't put in as much effort as (s)he should
09 (b)	Not involved in lessons	Positive approach
10	Withdrawn	Outgoing
11	-	-
12	Quiet	Vocal

In the construct from Teacher 07G we see an example of a distinction which was clearly in line with the test based measure of potential. The pupils labelled as high ability but $\bar{P}$ may have been a pupil about whom the teacher changed his mind, or may indeed have been a borderline case for nomination as a more able pupil when that judgement was being made.

Construct 09 (a) is perhaps an interesting example of a case where the teacher misinterpreted the relationship linking the distinction which he made between the pupils in the triad which generated this construct, and high potential. One can speculate that this teacher was looking for extensive effort as an indicator of high ability and undervalued the pupils who merely 'jogged along'. This would indeed be confirmed by a separate but similar construct elicited from this same teacher and reported in Table 11.3.

It was also suggested above, that the left hand poles of these constructs might be used rather like disqualifying cues. It is interesting, therefore, to see the similarities between many of the characteristics in this table and the relevant ones in Tables 11.5 and 11.6.

In Table 11.8 we give the constructs elicited from Type F1 and F2 Triads.

TABLE 11.8

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT $\overline{\text{TP}}$ pole	 TP pole
03	Poor understanding of concepts	Good understanding of concepts
04H	Quick to grasp ideas	Slow to grasp ideas
04M	-	-
05	Mature in speech	Doesn't use verbal communication effectively
07B	Keen	Not keen
07G	High level of performance	Low performance
08	Neat presentation	Untidy presentation
09 (a)	Neat written work	Untidy written work
09 (b)	Not involved in lessons - few questions. ideas, suggestions	Positive approach - plenty of questions, ideas, suggestions
10	-	-
11	-	-
12	-	-

Constructs 03 and 09b are perhaps examples of constructs where the teacher places too little emphasis on the left hand poles as indicators of low potential. The left hand poles of most of the remaining constructs could reasonably be interpreted as 'favourite cues'. In Constructs 04H and 07G the nature of the descriptions which appear under the right hand pole seem to be difficult to reconcile with the fact that the pupils described were regarded as more able both by the test and the teacher. This is discussed more fully in Note 7 at the end of this chapter.

Finally in Table 11.9 we give the constructs from random triads and from critical triads that were split in an unexpected way.

TABLE 11.9

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT *A PRIORI* CLASSIFICATION

TEACHER	CONSTRUCT		
03 (a)	Lacking in confidence	-	Plenty of confidence
03 (b)	Spontaneous interest in physics	-	Lacks spontaneous interest
03 (c)	Poor practical skills	-	Good practical skills
04H(a)	Asks a lot of questions	-	Asks few questions
04H(b)	Very interested	-	Shows little interest
04M(a)	Tendency to be silly	-	Sensible attitude
04M(b)	Does not work hard	-	Hard working
04M(c)	Prepared to think about physics problems	-	Wants to be spoon fed
05 (a)	Independant worker	-	Needs to work as part of a group
05 (b)	Uses ability	-	Tends to sit back
05 (c)	Reluctant to put forward ideas	-	Contributed to lessons
05 (d)	Shows ability in classwork	-	Quiet in classwork
05 (e)	Willing to take part	-	Not actively involved in work
05 (f)	Able in subject	-	Not able in subject
05 (g)	Works unnecessarily hard	-	Knows when to stop
07B(a)	Difficult to motivate	-	Enthusiastic
07B(b)	Extrovert	-	Not outgoing
07B(c)	Wants to succeed	-	No drive to succeed
07B(d)	Poor presentation	-	Good presentation
07G(a)	Underperforms relative to ability	-	Performs to level of ability
07G(b)	Working class aspirations	-	Middle class aspirations
07G(c)	Quiet retiring personalisty	-	Forward personality
07G(d)	Minimum parental support	-	Extensive parental support
07G(e)	Clerically efficient	-	Not clerically efficient
07G(f)	Highly motivated	-	Lower motivation

Table 11.9 is continued.

Table 11.9 continued.

TEACHER	CONSTRUCT
07G(g)	Needs high degree of personal supervision - Needs only minimum supervision
07G(h)	Exceptional interest in physics - Not interested
08 (a)	Not precise in explanation - Precise in explanation
08 (b)	Underperforms - Performs as well as able
08 (c)	Lazy - Conscientious
08 (d)	Low performance - High performance
08 (e)	Overestimates own potential - Realistic estimate of own potential
09 (a)	Takes an active part in lessons - Diffident in lessons
09 (b)	Restless - Normal conforming behaviour
09 (c)	Very competent - Poor performance
09 (d)	Conforms to school policy - Does not conform to school policy
09 (e)	Devoted to subject - Not devoted to subject
10 (a)	High overall intelligence - Low overall intelligence
10 (b)	Displays leadership qualities - Doesn't display leadership
10 (c)	Will actively seek help - Does not actively seek help
10 (d)	Asks questions freely - Rarely asks questions
10 (e)	Test performance above classroom performance - Balanced performance
10 (f)	Doesn't understand physics concepts - Does understand physics concepts
10 (g)	Positive attitude to subject - Negative attitude
10 (h)	Confident in his/her ability - Not confident in ability

Table 11.9 is continued.

Table 11.9 continued.

TEACHER	CONSTRUCT
11 (a)	Introvert - Outgoing
11 (b)	Lazy - Work oriented
11 (c)	Antagonistic to physics - Keen towards physics
11 (d)	Outspoken - Quiet
11 (e)	Competitive against others - Compete with own standards
11 (f)	Doesn't do homework - Does homework
11 (g)	Stands back from practicals - Gets on with practicals
11 (h)	Criticism has little effect - Criticism has effect
11 (i)	Not imaginative - Imaginative
11 (j)	Does best without being pushed - Needs to be pushed to do best
11 (k)	Very articulate - Not articulate
12 (a)	Reluctant to get on with work - Gets on with work
12 (b)	Likely to answer questions - Unlikely to answer questions
12 (c)	Haphazard in work - Methodical in work
12 (d)	Confident at subject - Lacks confidence at subject
12 (e)	Interested in physics out of school - Not interested out of school

This table completes the list of constructs elicited in our study of physics teachers, and allows us to make a general comparison between our overall results and those obtained by Wood and Napthali (1975) in their study of the constructs that teachers of geography and physics used in the assessment of the achievements of secondary school pupils. All but two of the eighteen constructs which they listed as having been mentioned by three or more of their teachers, are to be found in Tables 11.3 to 11.9 above. Most of their other constructs are either matched exactly in our findings, or find close analogies there. There would seem to be little indication, at least in comparison with that source, that major areas of our teachers' construct systems, which had relevance to their assessment of potential, were overlooked. We return to this point later.

It is also interesting to note that Wood and Napthali found that 'cognitive constructs play a central, although not overwhelming part, in the assessment of achievement', and that 'motivational' constructs could have a moderating effect on these assessments.

The results we have presented above would certainly seem to support the idea that cognitive and motivational constructs each have roles to play in forming teachers' overall opinions of pupils ability. By demonstrating that motivational constructs arise particularly from triads of Types B-F, our work would seem to suggest that teachers may be led to over- or under-estimate potential because of high or low motivation on the part of the pupil. We will return to this point later.

11.2 Preliminary conclusions in Physics

The over-riding conclusion from the work reported above is that physics teachers made use of fairly general concepts to do with ability in physics, and, of a wide range of motivational and personality based constructs.

It seems that they did not use constructs relating to the kinds of specific scientific behaviours that were mentioned on the checklists. This was even true of Teachers 03, 08 and 09 who had used the checklist in their identification.

One could suggest that the checklist had not made a lasting impression on teachers in terms of the way in which they looked at their pupils.

There is also an indication, mainly from Tables 11.4 and 11.5 that 'model pupil behaviour' (e.g. enthusiasm, confidence, conforming behaviour, concentration) acted as a favourite cue for teachers, and was associated with over-estimations of pupils' potential; and that qualities such as disorganisation and carelessness or quiet, withdrawn behaviour acted as disqualifying cues and were associated with under-estimation of potential.

11.3 Further analysis in Physics

Before presenting similar results in the other subjects we will pause to look a little more deeply at three matters concerning the physics constructs.

i) Our first concern is one that was raised in Chapter Ten: namely whether evidence can be found from the teachers' ratings of pupils to support the hypotheses that we have put forward that some constructs might have poles that were acting as favourite or disqualifying cues.

These hypotheses emerged from a consideration of the nature of the triads from which the constructs were elicited. If a teacher distinguished an under-estimated ($T\bar{P}$) pupil from two accurately judged (TP) pupils, one can certainly say that the pole of the resulting construct which described the $T\bar{P}$ pupil was associated with under-estimation of ability, and may have acted rather like a disqualifying cue, for that individual pupil. What cannot be said with certainty is that the teacher used this same cue in the same way in forming his/her judgement about other pupils.

If this could be established we could say that the cue is consistently associated with under-estimation on the part of the teacher - and that it would be worth attention as a possible general cause of under-estimation in the judgements made by that teacher. If it cannot be established, we can only make the weaker point that the cue was associated with under-estimation in the case of the individual pupil involved in the triad - and that it may be worth attention as something which may lead the teacher to errors of under-estimation on some occasions.

By looking at the way in which the teacher rated a number of pupils on the construct in question, and by comparing these ratings with the teacher's overall judgement of potential, we can begin to get some information that may help to clarify this point. We will discuss three examples, (a) - (c), below. In each case we will discuss pupils from the A or B Groups (i.e. pupils who had a test score which was high enough to imply that they were all good, or reasonable candidates for selection as more able). In making choices between these pupils, factors which affect the teachers' judgement are particularly likely to show up as the teacher is not in a position to respond simply to large differences in pupils' ability.

Example (a) Teacher 04H: a possible disqualifying cue.

In Table 11.5 'Slow to grasp ideas' was identified as a possible disqualifying cue for Teacher 04H. Table 11.10 shows how that teacher rated all of the A and B Group pupils whom he taught, on the construct containing that cue. We have divided the pupils into those whom he judged to be more able, and those whom he decided should not be in that category.

TABLE 11.10

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON
CONSTRUCT 04H(b) FROM TABLE 11.5

	Slow to grasp ideas			Quick to grasp ideas	
	1	2	3	4	5
SELECTED	0	0	0	2	4
NON-SELECTED	2	1	2	4	1

Some comments are made on the statistical significance of the association between selection and rating on the construct in Note 8 at the end of this chapter. The note also comments, in similar fashion on Tables 11.11 to 11.13.

Although the numbers in this table are small, we see that 5 out of 5 pupils who were rated 1,2 or 3 (i.e. at or near the proposed 'disqualifying cue' pole) were denied membership of the top 10% group, despite the fact that these pupils, like all of the pupils in the table, had a reasonable test score. This confirms the idea that 'Slow to grasp ideas' may have acted generally as a disqualifying cue for that teacher. Interestingly a rating of 5 at the other end of the construct also seemed to distinguish between selected and non-selected pupils. The pole 'Quick to grasp ideas' seemed to operate as a favourite cue. Although on a priori grounds we have made no case that it should do so from the appearance of the construct in Table 11.5, the same construct does appear in Table 11.6 where we would wish to suggest that this role of favourite cue might be expected.

Example (b) Teacher 04H: a possible favourite cue.

Table 11.11 gives similar data in connection with a construct from Table 11.4 in which one pole, 'Confident in practical work', appeared to operate as a favourite cue.

TABLE 11.11

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON
CONSTRUCT 04H FROM TABLE 11.4

	Hesitant in practical			Confident in practical	
	1	2	3	4	5
SELECTED	0	0	0	2	4
NON-SELECTED	1	1	6	1	1

In this example 6 out of 8 pupils who were rated 4 or 5 (i.e. at or near the proposed 'favourite' pole) were selected as more able by the teacher. It does seem that the pupil characteristic 'Confidence in Practical Work' was generally associated with teacher selection and may have been involved in tipping the balance in favour of a number of pupils when the teacher came to make up his list.

Example (c) Teacher 08: Possible disqualifying and favourite cues.

In Table 11.6 we suggested that the left hand poles may be favourite cues, and that the right hand poles of the same constructs may be disqualifying cues. Table 11.12 shows how Teacher 08 rated A and B Group pupils on one construct of this kind: 'Committed to the subject - not committed to the subject'.

TABLE 11.12

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON
CONSTRUCT 08 FROM TABLE 11.6

	Committed to subject			Not committed to subject	
	1	2	3	4	5
SELECTED	4	2	0	0	0
NON-SELECTED	0	1	3	1	1

6 out of 7 pupils rated 1 or 2 (at or near the proposed 'favourite' pole) were selected as more able by the teacher, and 5 out of 5 rated 3, 4 or 5 (at or near the 'disqualifying' pole) were not selected. This confirms the idea that 'Committment' operated as a favourite cue for this teacher in the assessment of the whole group of pupils, and that 'Non-Committment' operated, also in a general way, as a disqualifying cue.

Sometimes the tables needed to carry out these checks on a construct could not be constructed, either because very few A or B Group pupils were overlooked by the teacher (this was true of Teacher 03 and made it impossible to look at any of his constructs in this way), or because the range of ratings that was given to pupils was very small (this was true of the construct from Teacher 05 in Table 11.4). However the implication of these examples is that it is reasonable to suggest that the cues that were identified as disqualifying, or favourite cues on the basis of the nature of the triad from which they came, may well have operated in those same roles, in a consistent fashion, when the individual teacher went on to judge pupils who were not part of that triad.

All of the discussion above has related to further study of individual teachers. In the same vein it would be interesting, and potentially even more useful, to devise some further experimental test of whether the generalisations that we have made in discussing the nature of favourite and disqualifying cues (e.g. that 'model pupil behaviour' may act as a favourite cue for many of the teachers in our sample) are applicable to a large number of teachers. Such a study might be undertaken by asking teachers to judge whether pupils described in short pen sketches were able or not. The 'model pupil behaviour' element of the pen sketch would be varied systematically and an attempt could be made to see whether this variation had the effect that would be expected if this component of the description really did operate as a favourite cue for the whole group of teachers. We hope to be able to undertake such a study in the future, but can do no more at this stage than to suggest that it might be of value.

ii) We turn, now, to the second point that we wish to begin to explore here.

It may be useful to consider a classification of the constructs which have arisen in the study and to determine whether the constructs elicited from Type A triads (constructs associated with teachers' judgements which are in line with the pupils' test scores), those coming from triads of Types B to F (constructs associated with over- or under-estimation of potential) and those coming from random triads were of similar, or of different kinds.

Clearly such a classification should reflect the kinds of constructs which were elicited. However, in addition, there should also be some theoretical basis for each of the categories in the classification. Since we are interested in the relationship between teachers' constructs and pupils' future performance, we have worked on the basis that the classification should categorise constructs according to the kinds of factor which may be expected to influence pupil performance. Ability factors are clear candidates for inclusion. Youngman (1979) is one of many researchers who have shown that ability measures have considerable power as predictors of future performance. This has been confirmed in that part of our own work which is described in Appendix 1. In Chapters One and Two, and in Appendix 1, wehave discussed the role of specific and general aspects of ability. This discussion provides theoretical support for the inclusion of both specific ability and general ability components in the framework. The choice of physics ability and mathematics ability as the two sub sections of the specific ability component reflects the nature of the specific ability constructs that were elicited from teachers.

A considerable amount of research suggests that non-ability factors can also affect performance. Cohen and Manion (1981) point out that personality factors (especially extroversion/introversion) can have a considerable effect on performance. In addition, Entwistle (1968) established a clear relationship between motivation and achievement. It was therefore important to include these components in the framework. However, as colleagues who attempted to classify the constructs which we elicited, had some difficulty in agreeing whether some of these constructs were motivational or personality constructs, it was decided to combine these two categories in the final version of the framework that is set out in Figure 11.1.

Other influences on performance can also be traced in the literature. Brookover et al (1967) showed that pupil's academic self-image was correlated with grade point average. However it may be difficult for a teacher to be aware of a pupil's own opinion of his/her ability, and certainly no constructs relating to academic self-image were elicited. This component was therefore omitted from the framework.

Finally a wide range of situational factors such as teaching style (Eggleston et al, 1976), ability grouping (Postlethwaite and Denton, 1978) and general school ethos (Rutter, 1979) have been shown to relate to pupil performance to varying degree. However for a given teacher, in a given school, there was no variation in these factors across the pupils whom he/she was attempting to assess. As a result there were no constructs related to these factors, and they were therefore omitted from the classification framework.

Figure 11.1 summarises the classification which we designed in the light of these considerations.

FIGURE 11.1

A Framework for Construct Classification

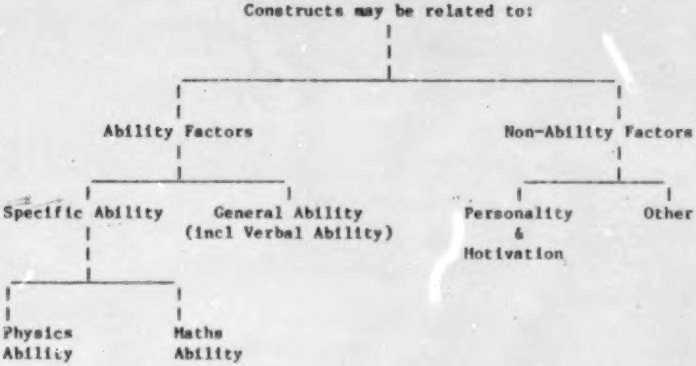

Constructs may be related to:
Ability Factors
Non-Ability Factors
Specific Ability
General Ability (incl Verbal Ability)
Personality & Motivation
Other
Physics Ability
Maths Ability

Table 11.13 shows the number of constructs of each type which fell into each part of the classification framework.

TABLE 11.13

THE NUMBER OF CONSTRUCTS OF EACH TYPE IN EACH CLASSIFICATION CATEGORY

Triad Type	Physics Ability	Maths. Ability	General Ability	Personality & Motivation	Other
A	4	2	4	3	0
B-F	6	1	14	32	2
Random	1	1	4	47	7

It is not surprising to see, from this table, that teachers were sensitive to factors such as the personality and motivation of their pupils. As indicated in the discussion above, these factors have a direct influence on the quality of the work done by pupils. However, such factors also colour the general relationship between the teacher and the pupil, and may be given too much weight when the teacher tries to predict the pupil's future performance.

It is interesting that the Type A triads which, on *a priori* grounds were likely to give rise to constructs which were most likely to help the teacher to make judgements about pupils that were in line with the tests, were largely ability constructs. On the other hand, constructs from the triads of Types B to F (which we expect to be associated with judgements that were not in line with the tests) were largely personality, or motivational constructs.

The implication is that though motivational constructs may have some predictive power as indicators of future potential (a fact that teachers seemed to recognise as is indicated by the presence of some motivational constructs amongst those elicited from Type A triads), physics teachers did indeed seem to place too much weight on them in forming their overall view of pupils' potential; (as is indicated by the large number of motivational constructs amongst those elicited from Type B-F triads).

Teachers might therefore make judgements which more closely matched with the test results if they were to attempt to make a clearer distinction between their assessments of pupils on motivational/personality dimensions and their assessments on ability dimensions, and if, subsequently, they were to place less weight on the motivational/personality dimension in coming to their overall judgement. To this end the detailed nature of the motivational/personality constructs that we have noted in Tables 11.4 to 11.8 may be helpful in alerting teachers to the kinds of motivational/personality information which might be affecting their view of pupils. Of course, not only can we draw teachers' attention to this overemphasis on motivational constructs we can also, through the classification of characteristics as disqualifying or favourite ones, indicate the way in which these constructs were being misinterpreted. This detail may help teachers to modify their response to such characteristics in such a way as to bring their overall judgement of potential more closely in tune with the pupils' test results. It may be helpful at this point to remind the reader that a closer match between teacher and test judgement is a worthwhile goal in physics. As shown in Chapter Three, the extent of the mismatch between teacher- and test-based judgement is too great to be accounted for by test error alone. In bringing the teachers' judgement more closely in line with the test judgement we would argue that teachers would be improving their assessment of pupils' potential for future performance in the subject.

One further point arises from Table 11.13 and that is that general ability constructs were more apparent from triads of Types B to F than from Type A triads. By analogy with the argument above, this would suggest that teachers should try to distinguish more clearly between specific and general abilities and, subsequently, to place more weight on the specific ability judgements when forming their overall view of pupils' potential. This point is in complete accord with points made elsewhere in this report where overemphasis on general ability on the part of physics teachers has been noted.

iii) The third, and final area in which we would seek to make more detailed comment concerns the completeness of the construct data.

We mentioned in Chapter Ten that teachers may go about the process of forming an overall judgement of their pupils by combining the data which they have for a pupil on each individual construct by what amounts to a multiple regression procedure - albeit probably an unconscious one. We applied such a model to our construct data and generated, for each teacher, a combination of the individual construct scales that provided the best estimate of the teacher's overall judgement on his/her pupils. If we had a reasonably full map of the teacher's construct system, (and if teachers did actually use some procedure which was analagous to multiple regression), then we expect to be able to explain a fairly high percentage of the variance in the teacher's overall judgement by means of the composite score from the combination of constructs. Table 11.14 gives a measure of the variance which was actually explained in this way.

TABLE 11.14

THE MULTIPLE REGRESSION MODEL APPLIED
TO TEACHERS' OVERALL JUDGEMENT OF POTENTIAL

Teacher	% Variance explained	no. of pupils	no of variables included in equation
03	72	27	3
04H	43	30	1
04M	31	33	1
05	45/23 (*)	28/28	1/1
07B	33	26	1
07G	23	26	1
08	62	28	2
09	40/61 (*)	22/24	1/1
10	19	29	1
11	60	33	1
12	57	37	1

(*) Teachers judgements on two separate classes were analysed separately.

In a large proportion of the cases in this table, a considerable amount of the variance in teachers' overall judgement could be explained on the basis of the construct ratings. For example in 61% of cases, 40% or more of the variance could be explained. There would therefore seem to be some evidence for the view that our map of the factors which influence teachers was complete enough to warrant the kind of detailed discussion which we have engaged in so far in this chapter.

11.4 Basic results in English

Tables 11.15 to 11.21 show the constructs which were elicited from English teachers. A short comment follows each table.

TABLE 11.15

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT	
	TP pole	$\overline{\text{TP}}$ pole
04A	Produces work of depth	- Does not produce work of depth.
04D(a)	Syntactically accurate	- Chaotic sentence construction.
04D(b)	Inventive writing	- Reliable, but just steady writing.
05C(a)	All round intellectual ability.	- Weak all round.
05C(b)	Exact vocabulary	- Imprecise vocabulary.
05R(a)	Not responsive	- Easy to deal with.
05R(b)	Clever, good background	- Less mature.
11A(a)	Generally able	- Generally weak.
11A(b)	Confident reading	- Hesitant reading.
11A(c)	Accurate work	- Inaccurate work
11H	Good O-level candidate	- Weak O-level candidate.

The main feature of these constructs (which were associated with teachers' judgements that were in line with the test results) was that they indicated sensitivity, on the part of the group of English teachers, to a number of apparently distinct English skills. This can be contrasted with comparable constructs in physics, where the only specific skill that was mentioned was that of mathematics ability. In common with physics teachers, English teachers also made reference to general ability in this table.

TABLE 11.16

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	CONSTRUCT		
	$\bar{T}P$ pole		$\overline{TP}$ pole
04A(a)	Confident	-	Not confident .
04A(b)	Not aware of what goes on in English.	-	Aware of what goes on in English.
04D	-		-
05C(a)	Strives for eloquence in emotional register.	-	Imprecise and inelegant.
05C(b)	Frequent signs of brilliance	-	Sound.
05R(a)	Thoughful and imaginative	-	Not evidently sensitive
05R(b)	Co-operative	-	Awkward in class.
11A	-		-
11H	-		-

We would remind the reader that the description at the left hand pole of these constructs may be acting as favourite cues for the teacher, leading him or her to nominate pupils for membership of the more able group who, in test terms, were unlikely to qualify for membership. In physics these favourite cues were largely to do with something that we labelled 'model pupil behaviour'. There are echoes of this in the English results (04A(a), 05R(b)). However one 'favourite cue' in English ('Strives for eloquence in emotional register') may be interpreted as an example of a teacher responding to an appropriate characteristic of high ability in English. This could have led to overestimation of the pupil in comparison with the test simply because the test was not able to measure such a sophisticated aspect of English ability. The same comment might apply to the favourite cue in construct 05R(a).

TABLE 11.17

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT	
	$\overline{TP}$ pole	TP pole
04A(a)	Does not take work far enough	Takes trouble .
04A(b)	Not original in writing	Produces original writing
04D(a)	Unwilling to commit self to an opinion	Willing to commit self to an opinion
04D(b)	Does not contribute orally	Expresses opinions and asks questions
05C(a)	Shows moral sensitivity & is idealistic	Has accurate insights into the wider world
05C(b)	Carries own social world with him	Excited about social possibilities
05R(a)	Less forceful	Forceful
05R(b)	Not keen	Keen
11A	Poor vocabulary	Good vocabulary
11H	Poor vocabulary	Has a feel for words/Good

In two places in this table there are echoes of the kinds of attitudinal disqualifying cue that we found in physics (04A(a) and 05R(a)). However we can also see, again, evidence of a sensitivity, on the part of teachers, to subject-specific skills that may have been overlooked by the test (04A(b), 11A, 11H). We suggest that, under the $\overline{TP}$ pole of these constructs, we find things which the test did not measure, and which the pupil did not do well. The teachers may have been influenced by these things in coming to their decisions that pupils were not more able. Since these subject specific weaknesses on the part of the pupils were not revealed by the test, the teacher and test disagreed. However the implication from the nature of these descriptions, is that it may be the test-based prediction which was in error.

TABLE 11.18

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT	
	$\bar{T}P$ pole	$T\bar{P}$ pole
04A(a)	Mature work	- Immature work
04A(b)	Normally popular	- Not popular.
04D	Good technical skills	- Weak technical skills.
05C	Impassioned in use of language.	- Steady in use of language.
05R(a)	Enthusiastic	- Not forward.
05R(b)	More disorganised	- Takes a lot of trouble.
11A	Gets comprehension exercises right	- Gets comprehension exercises wrong.
11H	-	-

The constructs in this table seem to be consistent with those elicited in Tables 11.16 and 11.17. Most of the favourite cues (at the $\bar{T}P$ pole) and disqualifying cues (at the $T\bar{P}$ pole) that are found here can also be found, in identical or similar form, in the appropriate earlier table.

TABLE 11.19

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT	
	$T\bar{P}$ pole	$\overline{TP}$ pole
04A(a)	Reliably good work	- Sometimes produces good work.
04A(b)	Great sense of humour.	- Reasonably quiet.
04D(a)	Good technical skills	- Poor technical skills.
04D(b)	Assured sophisticated style	- Underdeveloped style.
05C	Does not miss the point	- Intellectually weak.
05R	Confident	- Lacks self confidence.
11A	Has natural ability	- Has no natural ability.
11H(a)	Generally able (written work grammar,sentence construction)	- Not able in these ways.
11H(b)	Good O-level candidate	- Weak O-level candidate.

If one attempts to interpret the left hand pole of these constructs as disqualifying cues, one is led into some difficulty as most of them refer to high subject-specific ability. In this respect English and physics teachers differed, as most of the corresponding physics constructs (Table 11.7) mentioned poor attitudinal characteristics which did fit sensibly into the disqualifying cue model. The apparent inconsistency in English can be resolved by referring back to Table 10.1 where we first set out the interpretation of constructs arising from critical triads. In connection with Type E triads, we put forward an alternative interpretation. This was that the teacher may have placed too little weight, on the characteristics described under the $\overline{TP}$ pole, as indicators of high potential in the subject. From the nature of the constructs, this would seem to be the more appropriate interpretation in English.

TABLE 11.20

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT $\overline{TP}$ pole	TP pole
04A	Not aware of what is happening in lessons	Conscientious and studious.
04D(a)	Unimaginative	Imaginative.
04D(b)	Not committed to the subject	Committed to the subject.
05C	Satirical streak	Makes conventions work.
05R	Clever but quiet	Clever and not quiet.
11A	-	-
11H	Essays not sensitive	Sensitive Essays.

The $\overline{TP}$ pole of Construct 05C seems to be one which could be interpretted as a favourite cue. Other constructs (espcially 04D(a) and 11H) would seem to be things to which the teachers gave too little weight as indicators of low potential for the subject.

It may seem to be inconsistent to argue that the constructs arising in connection with triads of type B-F were, in physics, indicative of pupil characteristics which led teachers into errors of judgement, whereas, in English, they were largely indicative of appropriate characteristics which the teachers noticed, but which the tests did not measure. However, we would suggest that two things support this difference in interpretation. First, it must be remembered that, in Chapter Three, we found evidence that physics teachers did make errors of judgement in assessing pupils' potential for O-level, whereas we found no evidence to suggest that English teachers made such errors. It was possible to account for the full extent of the mismatch between English teachers judgements and English test scores entirely in terms of error in the test. Secondly the nature of the constructs themselves were consistent with the two different interpretations. Indeed one might argue that the consistency between the findings of Chapter Three, and the nature of the kinds of constructs found in English and physics was indirect evidence, suggestive of the validity of the critical triad methodolcgy.

TABLE 11.21

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT *A PRIORI* CLASSIFICATION

TEACHER	CONSTRUCT
04A(a)	Lively, enjoys contributing orally - Fairly passive in class
04A(b)	Mature - Immature
04A(c)	Nervous, anxious to do well - Confident
04A(d)	Generally more able - Generally less able
04A(e)	Depth in work - Superficial work
04A(f)	Unable to relate to what is going on - Integrated in to work of class
04A(g)	Poor written work technically - Technically competent in written work
04D(a)	Careful fastidious worker - Driven by inspiration
04D(b)	Mature - grown up for age - Not grown up for age
04D(c)	Makes witty remarks - Makes few remarks
04D(d)	Does not read books - Usually has a book on the go
04D(e)	Lazy/slipshod - Works with care and accuracy
04D(f)	Tries to incorporate new into work - Does not incorporate new ideas
04D(g)	Little improvement in quality of work - Improvement in quality of work
04D(h)	Sullen - Cheerful
04D(i)	Not good at listening to instructions or advice - Takes note of instructions and advice
05R(a)	Articulate - Reserved
05R(b)	Noticeable personality - Retiring personality
05R(c)	Confident - Quiet
05R(d)	Gives in class - Does not contribute in class
05R(e)	Chatty in class - Restrained in class
05R(f)	Responsive to English - Amenable but not responsive
05R(g)	Diffident - Not diffident
05R(h)	Supportive to teacher - Regular member of class
05R(i)	Does not take part in class - Part of lessons

Table 11.21 is continued.

Table 11.21 continued

TEACHER	CONSTRUCT	
05C(a)	'Meaty' use of language	Stylistic use of language
05C(b)	Little talent	Talented orally/practically
05C(c)	Poor stamina for work	Has stamina
05C(d)	Exhibits deftness and control	English is an uphill struggle
05C(e)	Perceptive, imaginative	Not receptive or imaginative
05C(f)	Doesn't understand English tasks	Does understand English tasks
05C(g)	Reliably sound	Assertively bright/sensitive
05C(h)	Sober seriousness	Not guided by predetermined standards
05C(i)	Light	Strives for depth
10T(a)	Quiet	Fluent orally
10T(b)	Excellent reading aloud	Poor reading aloud
10T(c)	Consistent top stream performance	Late developer
10T(d)	Quick thinking - work comes easily	Ponderous worker
10T(e)	Idle	Hardworking
10T(f)	Needs to be drawn	contributes without being drawn
10T(g)	Outstanding vocabulary	Limited vocabulary
10T(h)	Excellent oral work	Little oral work
10T(i)	Mechanically correct	Makes mechanical errors
10T(j)	Flair in written work	Mundane written work
10T(k)	Lacks confidence in work	Confident
10T(l)	Not a reader	A great reader
10T(m)	Sensitive artistic expression	Ordinary expression
10T(n)	Accurate description	Inaccurate description
10T(o)	Highly articulate	Not articulate
10T(p)	Writes beautiful poetry	Writes laboured poetry
10T(q)	Concise clear expression	Expression not concise and clear

Table 11.21 is continued.

Table 11.21 continued.

TEACHER	CONSTRUCT
11H(a)	Fluent writing - more ideas - Not fluent
11H(b)	Sensitive, expressive writing- Does not write sensitively
11H(c)	Poor comprehension - Good comprehension
11H(d)	Makes basic mistakes - Accurate
11H(e)	Good vocabulary - Poor vocabulary
11H(f)	Good grammar - Inaccurate grammar
11H(g)	Good ideas for essays - Weak ideas for essays
11H(h)	Lively mind - Not a lively mind
11A(a)	No interest - Interested
11A(b)	Poor attention - Attentive
11A(c)	Poor written work - Good written work
11A(d)	High accuracy in spelling and punctuation - Inaccurate
11A(e)	Tries hard - Does not try hard
11A(f)	Takes active part in lessons - Not involved in lessons
11A(g)	Lacks maturity of thought - Mature thinker
11A(h)	Gives up with deeper work - Perseveres
11A(i)	Generally able - Lacks ability
11A(j)	Has flair for use of words - Shows no flair
11A(k)	Has interesting style - Ordinary style

* This teacher's nominations matched test nominations so closely that no critical triads could be constructed.

11.5 General conclusions for English.

The main conclusions to be drawn from the English work are to be found in the brief discussions following each table of constructs. However it is interesting to offer some general, concluding remarks. We note that some 64% of English teachers' constructs were concerned with pupils' ability (48% being concerned with specific English ability). This can be compared with physics where only 29% of teachers' constructs were related to pupils' ability (and only 11% were related to specific science ability). In view of the greater accuracy of English teachers in identifying more able pupils, this is an important general point.

A subjective view of English teachers also emerged in the course of the elicitation interviews. We found that they were quick to respond with a considerable amount of information about individual pupils, and that this was often expressed with enthusiasm and genuine concern. For example, the teacher who described a pupil as writing 'beautiful' poetry really did mean the word beautiful and had incorporated some of this pupil's work into her own collection of poetry to be read for pleasure. It is encouraging that such involvement in the work of pupils is possible and interesting that this group of teachers proved to be so successful in their overall judgements.

11.6 Basic findings in French.

We now present the results of the study of the constructs used by French teachers. Unfortunately it was only possible to interview four teachers of this subject so the insights gained cannot be as full as they have been in physics and English. The main value of the work in French will be in indicating the kinds of constructs that teachers use. Generalisations about the nature of favourite and disqualifying cues will be difficult because of the small numbers of constructs of each type that can be studied. Nevertheless we will present the data in the same form so that it is easy for the reader to make some broad comparisons between the subjects should he wish to do so.

We give first the constructs arising from Type A triads.

TABLE 11.22

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT TP pole	$\overline{TP}$ pole
03	Highly able	- General lack of ability.
04J	-	-
04T(a)	Outstanding orally	- Lacks oral confidence
04T(b)	Generally competent	- Low general ability in French
05	Confident	- Not sure of self

It is interesting to see confidence constructs (05 and 04T(a)) in this table, (which, we have argued, is the table containing constructs which are appropriate to the judgement of potential in the subject), and to see the close link that is implied, in construct 04T(a), between confidence and ability. The results seems to some extent to be at odds with English and physics where the corresponding tables contained mainly ability constructs. The finding is, however, entirely in line with Pimsleur's (1964) view which we have quoted in Chapter Five. He stated that:

> ...the person who does well in a foreign language is one who is willing to get up and talk before others...

Other subject specific skills were not mentioned (in contrast to English where such skills were a major feature of the table). This does suggest that the English teachers' ability to notice these skills was not simply the result of the regular daily contact with particular groups of pupils (for the French teacher also enjoys this advantage in most schools). It argues that it is something in the nature of English teaching which enables the teachers to be aware of specific abilities in that subject.

In Tables 11.23 to 11.27 we give the constructs from the triads of Types B to F. In view of the very small numbers of constructs in each table, we will resist the temptation to draw further conclusions from each one.

TABLE 11.23

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	**CONSTRUCT** $\bar{T}P$ pole	 $\overline{TP}$ pole
03	-	-
04J	Generally more able in French-	Less able in French
04T	-	-
05	Sees patterns in French clearly -	Does not see patterns clearly.

TABLE 11.24

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT		
	$\overline{TP}$ pole		TP pole
03	Competent	-	Outstanding linguist
04J	Lacks imagination	-	Has imagination
04T(a)	Underperforms	-	High quality performance in subject
04T(b)	Poorly motivated to subject	-	Well motivated to subject
05(a)	Extrovert	-	Quiet academic
05(b)	Careful/fastidious	-	Slapdash

TABLE 11.25

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT		
	$\overline{TP}$ pole		$\overline{TP}$ pole
03	-		-
04J	Uses talent	-	Under-uses talent
04T	-		-
05(a)	Able	-	Not able or motivated
05(b)	Well motivated	-	Little committment

TABLE 11.26

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT $T\bar{P}$ pole		$\bar{T}\bar{P}$ pole
03	Serious about work in French	-	Underachieving
04J	-		-
04T	-		-
05	Brighter, more motivated	-	Afraid to let herself go

TABLE 11.27

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT $\bar{T}P$ pole		TP pole
03	-		-
04J	Has flair/natural ease with subject	-	Systematic/hardworking
04T	-		-
05(a)	Determined to make progress	-	Satisfied with own level of achievement
05(b)	Good at applying knowledge	-	Weak in applying knowledge

Finally, in Table 11.28 we give the constructs arising from random triads, and from critical triads that were not split in the expected way.

TABLE 11.28

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT A PRIORI CLASSIFICATION

TEACHER	CONSTRUCT
03 (a)	Highly competent - Has difficulties with technical details
03 (b)	Bad attitude to the subject - Enthusiastic for subject
03 (c)	Irrepressible gossip; talks not works - Able to concentrate
03 (d)	Quiet serious worker who achieves - Achieves little
03 (e)	Over pressured - At ease with work in French
03 (f)	Has confidence to participate in oral work - Lacks oral confidence
04J(a)	Lacks application - Has application
04J(b)	Unreceptive of advice from teacher - Keen to please
04J(c)	Gives impression of brightness - Steady
04J(d)	Alert - Vague
04J(e)	Impulsive - Painstaking
04J(f)	Seeks attention - Reclusive
04J(g)	Independant of mind - Amenable
04J(h)	Orally confident (extrovert) - No oral confidence (introvert)
04J(i)	Easily hurt - Resilient
04J(j)	Insouciant - Concerned to please
04T(a)	Confident - Not confident
04T(b)	Extremely quiet; gives little orally (or written) - Confident and lively
04T(c)	Willingly contributes - Contributes only if asked
04T(d)	Scruffy written work - Meticulous
04T(e)	Academically outstanding - Sloppy
04T(f)	High academic standard - Low academic standard
04T(g)	Extremely quiet - Positive
04T(h)	Sociable - Restrained
05 (a)	Highly able in languages - Lacks sense of purpose & effort
05 (b)	Co-operative - Awkward

11.7 General conclusions in French.

Confidence (particularly in relation to oral work) and motivation were clearly important components in the French teachers' maps of childrens' abilities. There were also several mentions of general 'Ability in French'. However, there were few references to subject specific skills, such as technical accuracy, memory or vocabulary, which we might have expected.

Some 35% of constructs were clear ability constructs, 10% related to confidence, 20% to motivation and 35% to clear non-ability constructs. The comparison between English and French teachers in terms of the kinds of constructs which they used is again of interest in view of the generally lower accuracy of French teachers in identifying the more able.

Perhaps the most important general conclusion is that, because of the small number of teachers who were involved, the work which is reported here should be considered as a pilot study of French teachers' construct systems. A larger follow up study might add much that is of interest - including insights into the nature of favourite and disqualifying cues used in this subject.

Constructs used in Mathematics.

As mentioned at the beginning of this chapter, a simpler study of mathematics teachers' constructs is provided in Appendix 8.

Notes for Chapter Eleven

1) It is interesting to note that one particular teacher (08) was responsible for four of the 12 unsplit critical triads.

2) Teacher 05 was no longer teaching in the same school when the interview was conducted and one would expect impressions of past pupils to be somewhat overlaid by new concerns. Teacher 11 was interviewed in a rather noisy staffroom.

3) F1 triads were the first type of critical triad to be presented and many teachers found two ways of splitting these triads. The splits for this triad type were not therefore independent, and this particular significance figure should be treated with caution.

4) The chance of making 8 expected choices from the 12 possible choices if the selection was based on a random split of the triad is

$$^{12}C_8 \cdot (\tfrac{1}{3})^8 \cdot (\tfrac{2}{3})^4$$

The significance level = $\sum_{n=8}^{12} \, ^{12}C_n \cdot (\tfrac{1}{3})^n \cdot (\tfrac{2}{3})^{(12-n)}$

5) The significance level associated with 6 expected choices out of 11 is 0.12 but that associated with 7 out of 11 is 0.04.

6) It may be helpful to remind the reader that the teachers did not know the pupils' test scores when they made their overall judgement of the pupils' potential, nor when they took part in the elicitation interview.

7) Two constructs from this table (04H and 07G) call for further comment as the descriptions found at the right hand poles of these constructs seem difficult to reconcile with the fact that the pupils who were being described were high scoring pupils who had been nominated as more able by their teacher (TP pupils). One is tempted to ask how it can be that a TP pupils can be described as 'Slow to grasp ideas' or as pupils who achieved only 'Low performance' levels. The solution to this problem in the case of Construct 04H is fairly straightforward. Within the triad, the two TP pupils were described as 'Slow to grasp ideas', in contrast

to the $\bar{\text{T}}$P pupil who was described as 'Quick to grasp ideas'. However when the teacher rated all of his pupils on this construct, all three who had formed the triad were rated as 'Quick'. One can interpret this as indicating that the $\bar{\text{T}}$P pupil was particularly quick, even in comparison with the two TP pupils, but that all three were quick in comparison with the rest of their classmates. This interpretation removes the puzzle over the TP pole, without calling into question the possibility that 'Quick to grasp ideas' was operating as a favourite one for that teacher.

Similar problems of this magnitude do not seem to have arisen in connection with most of the other constructs reported in this section.

We turn now to the apparent difficulty concerning the interpretation of the right hand pole of Construct 07G. One of the TP pupils who was described as showing 'low performance' can be accounted for by an argument similar to that set out above. His rating on the construct indicated that his performance was perceived as being lower than that of the $\bar{\text{T}}$P pupil in the triad, but it was still high compared with the rest of his class. For the other TP pupil in the triad the explanation of the apparent inconsistency is more complex. As can be seen from Construct 07G(a) in Table 11.9, this teacher attempted to distinguish between 'ability' and 'performance'. The second TP pupil who was described as showing low performance when the Construct 07G in Table 11.8 was elicited, was rated in such a way on Construct 07G(a) (Table 11.9) as to suggest that his ability was perceived as being higher than his level of performance. This higher ability on the part of the pupil was entirely consistent with his classification as TP.

8) The numbers in this table were too small to allow a chi-squared test to be carried out on the 5 by 2 table. Instead the table was regrouped as a 2 by 2 table (ratings 1 to 3 were combined, as were ratings 4 & 5). A Fisher Exact Probability test confirms that the new table showed a significant association between selection and rating ($p \leq 0.05$). An identical re-grouping of Table 11.11 was carried out and again a significant association was detected ($p \leq 0.005$). Table 11.12 was re-grouped by combining ratings of 1 & 2 and ratings of 3 - 5. This again showed a significant association ($p \leq 0.01$).

Table 11.13 was re-grouped into ability and non-ability factors to overcome small number problems. The resulting table showed a significant association between the triad type and the classification of the resulting construct ($X^2=27.3$, $df=2$, $p\leq0.001$). As can be seen from the table below the Type A triads gave rise to more ability constructs than expected, whereas the random triads gave rise to non-ability constructs.

Triad Type	Construct Classification Ability	Non-Ability
A	10 (3.75)	3 (9.25)
B-F	21 (15.89)	34 (39.11)
Random	6 (17.34)	54 (42.66)

Observed numbers are shown in the table with the relevant figures from the expected table in brackets.

If Type A triads were compared just with Type B-F triads, a significant difference between the two was found ($X^2= 4.89$, $p\leq0.05$). Again the Type A triads gave rise to more ability constructs and the Type B-F triads to more non-ability constructs.

CHAPTER TWELVE

Further Development and use of Checklists

12.1 Introduction

From the preliminary work with checklists it is clear that at the very least a different method of use must be employed if there is to be worthwhile effect on the identification of children of high ability in the classroom. The level of knowledge that a teacher has of each individual would seem to be insufficient for an adequate profile to be built up when the lists are filled in at one sitting. In other words it is not sufficient to merely draw the teacher's attention towards the appropriate characteristics for ability in the subject and away, say, from a markbook of accumulated evidence of performance, which gives at best a general overview of achievement over a period of time.

This is disturbing in itself because it shows that teachers may not have this important knowledge available. Even though, in such subjects as English and mathematics and less so in such subjects as physics and French, the general knowledge of subject specific ability may be adequate to predict within acceptable limits for success of O-level, the fuller knowledge of the profile of aptitudes is important if provision is to be tailored to the needs of the individual.

Furthermore, if a teacher cannot rely on his/her ability to gain detailed knowledge from day-to-day interaction with the pupils, he/she becomes underconfident in this ongoing method of monitoring progress, so that the benefits for the individual disappear. These benefits can be compared with the consequences of the once and for all method of administering a battery of psychological tests and the resulting labelling of children.

The conclusions that we have drawn indicate a problem, but we do not intend to make unjust accusations. The secondary teacher works under considerable constraints such as time, class size and syllabus, so that it is within these constraints that the identification process must proceed.

It may be possible to overcome these constraints to some extent by diverting more of the limited energies available, for the numerous tasks in the classroom, to attempt to improve the detailed identification guided by the checklists. It was with this in mind that we sought the co-operation of a sample of teachers in each subject to use the checklists in an appropriate way on new groups of third form pupils, during the second year of our research.

There were signs, from the first year, that the task was not completely hopeless, and that what was required were refinements to the method rather than dramatic change. The following analysis gave us some encouragement in this respect.

When we distributed the checklists for the first round, we asked teachers to fill them in for the top 20% of the year group and then to identify the top half of this group taking account pf checklist results. We also asked them to fill in the checklists for randomly selected pupils whose predicted O-level scores put them below the 80th percentile. In this way we could establish whether, even if the checklists could not separate abilities in the higher bands, they may still have enabled coarser identifications to be made. We scrutinised the completed checklists to try to discover if, in general, pupils ranked lower by tests were scored lower on the checklists. We would not wish to say, at this stage, that tables like the one below do any more than give us some encouragement that the broad separation of abilities is possible. We have not attempted to extract more from this data than is justifiable in the light of this point. For example we have not subjected the data to statistical tests. Nevertheless, we feel that there is some evidence that, in general, very weak pupils were scored lower than very good pupils. Table 12.1 gives figures for one school and subject. The distribution of percentages that is implied from this table illustrates what was to be found for each school and subject. The figures were derived from the number of checklist ticks in the Low, Medium and High categories, totalled for all checklist items. The results are given separately for pupils in each of the three bands: top 10% on checklist aided identification, next 10%, and random (below 80th percentile on the tests), and are expressed as percentages of the total number of checklist ticks.

TABLE 12.1

PERCENTAGE NOMINATIONS AS HIGH, MEDIUM OR LOW ABILITY ON CHECKLIST ITEMS, TOTALLED FOR ALL PUPILS (SCHOOL 10,MATHS)

	LOW	MEDIUM	HIGH
TOP 10% on checklist aided identification	0%	14%	86%
Next 10%	4%	56%	40%
Random below 80th percentile	57%	28%	15%

It may be that we are wrong to interpret these results as showing that checklists could, at least, be used for coarse separation of abilities. It could be that teachers have already labelled pupils into ability categories and merely fill in the checklists to reflect this labelling. This possibility is quite plausable if the detailed knowledge of the spectrum of abilities is not present even in the coarser sense. Nevertheless we lose nothing by treating the results optimistically and proceeding to our second phase of activity.

In the previous two chapters we have shown how a certain number of our sample of schools contributed to the other aspect of our study. This left a further restricted sample to co-operate with the development of checklists. The lack of difference between all schools in the first years work allowed us to work with these limited samples to discover general results that would apply to the wider sample.

The general strategy in English, French and physics was to negotiate the manner of checklist usage with each teacher, first to establish the relevant checklist items for the teacher's own broad aims and second to agree on the most appropriate method of use for the particular school environment. We were able to choose just one subject, because of manpower limitations, for more intensive classroom based activity. In all subjects, however, the general strategy was for the teachers to have the checklists at hand and in mind during the planning and teaching of lessons, so that the clues displayed, from time to time by the pupils, could be recorded and accumulated for a final scrutiny when it was considered that sufficient recording had taken place for valid identification of the top 10% to be undertaken.

Appendix 5 includes details of the developed checklists and recording forms.

We will now summarise the main points regarding checklist development for English, physics and French and then report on the analysis of results for these three subjects. We will then give a fuller report of the work done in mathematics before drawing final conclusions together.

12.2 Development of further use of checklists in English, French and Physics

ENGLISH

Discussions with teachers that used Year 1 checklists revealed no strong or consistent objections to individual items. In nearly every case, a small proportion of items (one or two) was identified as less satisfactory than the majority of items. Because of the fact that different English teachers, with their different approaches to the subject, invariably picked out different items for criticism suggested that no single item could be identified as inadequate. All items gained enough support to justify their inclusion in the second round checklists.

The chief lesson from Year One was the need for both a longer period of time in which to apply the checklists and a more simple and concise format. Thus, in conjunction with the other subjects, the English checklists were applied over one and a half terms. The format was redesigned in such a way as to eliminate the continuum approach, and a single description of each item was provided.

Each head of department who agreed to use the checklists was consulted over both the method of applying the lists, and over the choice of items. Where dissatisfaction over particular items was expressed, the teachers who would be using the lists were encouraged to modify them to suit their own criteria or means of assessment.

In every case, the departments involved entered into the process assured of the flexibility available to them in applying the lists, and the teachers involved appeared satisfied with the requirements of the task they had undertaken.

It transpired, as the term and a half elapsed, that the increased period of time did not - in most cases - lead to an increase in teachers' ability to use lists with confidence. Teachers who at the start had every intention of carefully checking the items over a long period of time still found it easier to deal with them in a more intensive way at some stage towards the end of the process. Almost without exception, they attributed this to the fact that they were being asked to apply criteria that - even if similar to those they might themselves select - had not been generated by themselves or (most crucially) by the work they had undertaken in that time.

This effect was evident consistently in all the schools (08, 09, 04, 00) using the checklists. The lists were filled-in in all but one case (half of school 09) but with little sense of satisfaction or accuracy in the opinion of the teachers. Most felt that previous knowledge and assessments of the individual pupils was applied onto the lists, rather than that the lists elicited new information about the pupils. Teachers felt that the exercise was artificial, in this form, although many did feel that the idea of a checklist was, in itself, valuable if it could be developed out of the syllabus that they had devised in their own department.

FRENCH

Although all the French teachers who used the checklists in Year One did succeed in completing them, there was a strong sense of dissatisfaction expressed over what they felt was a hurried and ineffective procedure. The short time-span seemed to accentuate problems with the items on the checklists: a lack of familiarity in the working of certain items emphasised the rather alien nature of this task. Unfortunately, as became apparent during discussion with teachers at this time, their existed very little middle ground between the kinds of item that elicit no new kind of insight about pupils (familiar concepts in familiar terminology) and the kinds of item that were meant to direct teachers to new insights about their pupils (unfamiliar concepts, or unfamiliar terminology) but turned out, in practice, to be more baffling than revealing.

The most evident fact at this stage was the limited number of criteria which French teachers felt were valid in assessing high ability in third years pupils. A need was clearly felt for a limited number of familiar and therefore clear items - but it was hard to see how a checklist designed along such lines could elicit fresh insights.

New checklists, with the least approved items discarded and certain new suggestions incorporated, were offered to teachers for Round Two. Teachers were encouraged to modify any individual items that they found excessively unsuitable, and were also encouraged to complete the lists steadily over the full one and a half terms.

Discussions with the four departments involved produced a fair consensus of approval for both items and the method proposed.

At two of the schools (04 and 08) the lists were used as planned and the items were greeted with a fair amount of approval, with the normal reservation that lists generated by the teachers themselves would be more effective.

At another two schools (09 and 00) the increased amount of time produced minimal success in achieving successfully completed lists. In both departments, the explanation was offered that the items did not relate, in practice, to the kind of work that was being carried out during this period.

Although the idea of the checklists met with general approval, it was obvious that there were many difficulties involved in using lists that were designed externally. One teacher summed up some of the difficulties as follows:

> One thing that comes out of the provision of the checklists is that we do not seem to have made very extensive use of them! Our intentions were no doubt always good, but unless we made a conscious effort at the time to take them out in class I certainly found that they were left incomplete for long stretches, and cannot be regarded as reliable indicators of anything more than observations made at irregular intervals.

PHYSICS

After they had completed the first version of the physics checklists, the teachers involved were consulted, so that ideas for the improvement of the lists could be accumulated. It transpired, from these discussions, that teachers were generally happy with the items on the checklist (though there were a very small number of instances where teachers were unsure about the exact meaning of individual items). However several aspects of the use of the lists caused problems for many of the teachers.

First teachers felt that it was very difficult to fill in the list reliably at one sitting. Some aspects of the pupils' scientific behaviour which the lists required teachers to assess, were only clearly displayed in connection with particular parts of the syllabus (indeed this was recognised in the design of the list by indicating syllabus areas where the teacher might expect to be able to observe the characteristics specified by the items.) Teachers felt very unsure of their ability to recall, sometimes after quite long periods of time, how each pupil matched up to each of the separate descriptions of scientific behaviour set out on the list. Several teachers mentioned that they were aware of the possibility of being influenced by a halo effect; so that a pupil known to perform in the way described by one item, would be assumed to be equally competent in terms of the other items.

Secondly teachers commented on a futher consequence of the fact that the opportunities for pupils to display some of the behaviours set out on the checklist were fairly rare. They felt that it was difficult to give a true picture of a pupil when the checklist asked them to comment on whether that pupil 'always, often rarely or never' behaved in accordance with the description of behaviour that was given on the checklist.

Finally some teachers felt uneasy about their ability to interpret the checklists once they had been completed.

Since there had been a little lack of clarity in the items themselves we decided, in preparation for the use of modified checklists in the second year of the project, to discuss each item with each of the teachers who were actually going to use the checklists. (As we pointed out in Chapter Five discussions with teachers had been an essential part of the design of the lists - but not all of the teachers who filled in checklists had been involved in those discussions.) These discussions were followed up by a summarising note and an example of such a note is included in Appendix 7. These discussions certainly clarified the thinking behind the items and also gave teachers the opportunity to reject individual items as being irrelevant to the particular kind of course which they teach in the third year. Teachers took very little advantage, in practice, of this opportunity to veto items.

To overcome the problems of completing the list at one sitting and of responding appropriately to the request to say whether pupils 'often, usually, rarely or never' displayed the high level behaviours indicated on the list we changed the format to that indicated below:

2nd YEAR PHYSICS CHECKLIST FORMAT

ITEM	PUPILS NAMES					
	Ann	Sarah	....	John	....	Simon
Selective in observing						
Seeks to explain not just describe phenomena						
....						

A full version of such a checklist grid is included in Appendix 7. Teachers were asked to use the checklist over a period of 1½-2 terms and to indicate, with a tick in the appropriate box, in a given child displayed one of the behaviours listed.

Although it would have been tempting to solve the problem of intepretation by asking teachers to count the number of ticks which accumulated for each child and to give, as their top 10%, those children with the most ticks we were still sceptical of the wisdom of using the lists in such an automatic way. As mentioned in Chapter Five, items varied in importance and, anyway, the checklist was intended to be used alongside other information which the teacher accumulated on his/her pupils. We resolved this problem by agreeing that the final selection of the top 10% pupils should be done by the teacher in consultation with one of the research officers. The procedure followed in most of the schools was that the teacher made tentative nominations for the top 10% list. The teacher and research officer then consulted the checklist and discussed the nominations taking particular care over pupils who seemed to be in the list of nominations but not to have many ticks on the checklist, and those who were not on the list of nominations despite a considerable number of ticks.

The teacher then made the final decisions and produced the top 10% list.

This procedure for completing and interpreting the lists was much more acceptable to the teachers - although, without exception, they did find it difficult to observe and record the scientific behaviour of their pupils in the midst of all of the other tasks which have to be undertaken during, and immediately after lessons. For some items few, if any ticks were recorded. This must throw into doubt the validity of the record provided by the fully completed checklists that were returned during the first year of the project, and may be one reason (perhaps the main reason) for the apparent ineffectiveness of the lists in improving teacher identification procedures on that occasion.

Effectiveness of the checklists

Before turning our attention to the more detailed work that was possible in mathematics we will look briefly at quantitative aspects of the second identification by checklists. The most appropriate table for this purpose is Table 12.2 which is rather like others from previous chapters, where we counted up how many identified pupils were in the categories 'good', 'satisfactory' and 'unlikely', based on their predicted O-level score. The reader may wish to refer to Section 3.4 to be reminded of the background.

We have compared the distribution of numbers of pupils who were nominated in these categories with the distribution of numbers identified from the first year's work without checklists. The comparison was by means of a one row chi-squared 'goodness of fit' test and was a way of establishing any quantitative change in the distribution of pupils over the three categories due to checklist usage.

The figures from the first year's work are extracted from Table 3.10.

TABLE 12.2

COMPARISON OF THE DISTRIBUTION OF PUPILS WHO WERE NOMINATED AS IN THE TOP 10% BY THEIR TEACHERS WHEN CHECKLISTS WERE USED WITH THE DISTRIBUTION WHEN CHECKLISTS WERE NOT USED

		NO. OF CHECKLIST AIDED NOMINATIONS	% OF NOMINATIONS WITHOUT CHECKLIST AID	EXPECTED NUMBERS FOR GOODNESS OF FIT TEST	X^2	SIGNIFICANCE
ENG	GOOD	35	61	31.11	3.48	NS
	SATISFACTORY	8	27	13.77		
	UNLIKELY	8	12	6.12		
FR	GOOD	23	51	24.48	1.41	NS
	SATISFACTORY	9	13	6.24		
	UNLIKELY	16	36	17.28		
PHYS	GOOD	26	45	18	6.54	5%
	SATISFACTORY	7	25	10		
	UNLIKELY	7	30	12		

We see that in English there was not a significantly different distribution of pupils in the three categories. We recall that in English there was no major concern from the preliminary work about the distribution of pupils in the three categories in schools that did not use checklists. The concern does still exist, however, that the teachers found difficulty in using the checklists so that their identifications via checklist were based, perhaps, on the same sort of knowledge as without them. So, although in English, their unguided assessments of pupils seemed very appropriate, the credit for this cannot, it seems, be attributed to checklist use.

Neither round of checklist usage has increased the level of French teachers' success in identifying the more able pupil. There is therefore the same indication, as in English, that the teachers were operating on the same kind of information whether they completed checklists or not. Of course the added problem in French is that earlier work, reported in Chapters Three and Eleven, suggested that this information did not seem to lead to accurate identification, nor did the component parts of it seem to be very appropriate.

In physics we see that significantly more pupils were chosen in the 'Good' category from this round, so that we are encouraged to believe that checklists can lead to improvement in the identification strategy, but we would not underestimate the effort that is required from the teacher who wishes to discover appropriate knowledge of the pupils' strengths and weaknesses in this way.

We can add to the above results by calculating the percentage of the high scoring A Group pupils who were identified from this checklist aided activity.

The percentages were: English 89%, French 61% and physics 83%. There would seem to be some concern in both physics and English that a number of these pupils were still missed, but a major concern in French that as many as 39% were missed.

12.3 Mathematics

Introduction

We will now describe the more detailed work that was possible in mathematics, starting with a description of the research aims and methodology.

One reason that checklists were not a very effective tool in the identification procedure could have been that the sorts of clues that are available to teachers both in the classroom and from scrutiny of pupils' workbooks, are inadequate. This could be associated with the inadequate challenges presented to the pupils in their work, so that the responses to these challenges would inevitably give the teacher inadequate information, or it could be that the environment in which the clues to ability occur will not allow adequate observations and records to be made.

The detailed knowledge of a pupils' profile of mathematical abilities was certainly absent when teachers were interviewed, yet the interviews did not provide an appropriate means of discovering reasons for this lack of adequate knowledge. There was, however, no criticism of the items on the checklist from the teachers, but there was a good deal of concern that the checklists were difficult to fill in, indeed there was some alarm from teachers that the constraints of time and syllabus did not allow the broad spectrum of abilities to be challenged. It was suspected that a reasonable O-level grade could be achieved by drill and exposure to appropriate questions, where 'tricks' could be learned that would lead to success. Because of the constraints under which the teaching proceeded this was a plausable route to O-level, rather than a strategy that aimed for breadth and from which O-level would naturally emerge.

The most appropriate research method, therefore, was to work closely with the teacher during this second phase of checklist usage and, by means of classroom observation, investigate the problems at first hand.

The first exercise was to develop the checklist into a better instrument. A careful re-assessment of the items was made with the co-operation of both teachers from the schools and tutors from the Oxford Department of Educational Studies.

The aim of the development work was to ensure that the checklist items were appropriate for observation at third form level as well as valid for the identification of pupils able in mathematics. A careful check on the validity of these items was carried out. This work is described fully in Appendix 7.

An adequate amount of time was agreed with each of the teachers who would co-operate with the work, so that one of the research officers would be allowed to observe lessons over that period. Both teacher and observer would record clues to ability as they arose in the course of the classroom interaction with pupils. At an appropriate time teacher and observer would accumulate the information on a master chart and attempt an identification of the top 10% of the year group. After this exercise was complete a reassessment would be made and action taken to improve the method if it proved inadequate. This would be followed by a second phase of observation. Two terms were allowed for the initial planning, one for each phase, but it was found to be possible to complete the work in two half terms.

The question of reliability of observations is important in work of this kind so proper measures were taken to include reliability checks. The observer used a special form for recording details of the clues to ability during classwork. The relationship of the checklist items to a sample of clues was then discussed with an experienced mathematics teacher. This teacher also observed a sample of the same lessons as the research officer and records of observations were compared. Examples of the observations are included in Appendix 7, as are details of this reliability work, so no further discussion will be included here, other than to confirm that reliability of the observations was impressively high.

The question of reliability connected with scrutiny of workbooks was tackled in a different manner. We will see that very little could be confidently recorded across the spectrum of mathematical abilities when pupils books were 'marked' in the conventional way. The reliability of the evidence that was assumed from this aspect of the work was checked in co-operation with a mathematics tutor at the University of Oxford Department of Educational Studies. He was asked to list the challenges that were presented by a sample of exercises from the textbooks that pupils used and this was checked against the research officer's assessment. A high level of agreement was reached.

Added to these strict checks of reliability were the ongoing discussions of a more informal nature between teacher and research officer concerning the sorts of clues that the work with the pupils had revealed. There was complete agreement in this co-operative aspect of the work, though it must be stressed that records of clues picked up by teacher and observer were recorded separately. This was in order that an assessment could be made of the effectiveness of the teacher's observation, within the constraints caused by other aspects of the teaching situation, compared with the possibilities that could be exploited by an observer with just that task in mind.

A second reliability check was made after the 10% lists had been finally drawn up. Random samples of pupils were chosen for intensive testing in a one to one situation so that their profile of aptitudes could be adequately challenged, levels recorded and compared with the conclusions from the observational work. This is reported in Appendix 7 and the reader will see that there was no conflict between conclusions from this work and what was concluded from the observations.

We will now discuss results of the observational work. These will be given for the groups of pupils, one by one, and then we will draw together the overall conclusions.

In the checklist tables that follow the checklist categories are coded. It is appropriate to remind the reader of these codes at this point.

CHECKLIST OF MATHEMATICAL APTITUDES

UNDERSTANDING OF INFORMATION

1. Able to understand the formal structure of a problem, seeing the relationship between the items of information even when there is a lot to assimilate.

2. Able to see when there's redundant information in a problem even when there is a lot to assimilate.

3. Able to see when there is not enough information in a problem and to see what information is needed for its solution even when there is a lot to assimilate.

PROCESSING OF INFORMATION

4. Able to suggest plausible ideas that might lead to the solution of a new sort of problem.

5. Able to follow the logic in a long mathematical proof.

6. Able to construct mathematical proofs:
   - a) similar to ones already demonstrated;
   - b) requiring different logic from that already demonstrated but on a familiar topic;
   - c) on problems related to a new or unfamiliar topic.

7. Quick to see when a method of solution that has been applied to problems in one topic will solve problems of a similar nature in another:
   - a) with help;
   - b) independently.

8. Quick to see when examples of a mathematical result have a wider and more general application: focuses readily on these generalisations rather than on the particular examples.

9. Able to follow a logical explanation or proof when valid logical jumps are made that miss out or assume some of the intermediate steps of logic.

10. Displays an ability to make valid logical jumps when using mathematical reasoning accompanied by the ability to provide the intermediate steps.

11. When a method fails in a process of reasoning is able to try an alternative approach incorporating what has already been validly reasoned into the new structure:
    a) in following someone else's logic;
    b) in constructing a system of logic independently.

12. Is able, when possible, to give alternative methods of solving a problem.

13. Is critical of inelegant solutions and strives in his/her own work for the most elegant.

14. Is quick to notice an incorrect step of logic or an invalid conclusion.

15. Is not content unless he/she has complete understanding of a mathematical concept, method or process of logic.

16. Adapts readily from a direct to a reverse train of thought.

MATHEMATICAL MEMORY

17. Has a good memory for a method of solution. Remembers a proof by its logic rather than by rote learning.

18. Has a good long-term memory for techniques developed for a particular topic. Requires little or no revision to regain previous levels of competance.

GENERAL

19. Shows little mental fatigue even after intense concentration on difficult problems.

20. Is not content to leave an interesting problem unsolved, but keeps returning to it at odd moments until success is achieved.

21. Can perform any arithmetical computation accurately and with ease.

22. When attempting to solve new sorts of problems displays a strong tendency to interpret and solve the problem in:
    - a) a spatial/geometric way; or
    - b) an analytic/logical way.

For the purposes of identification these were recoded according to the aptitudes that emerged from the work of V.A. Krutetskii (1976). The link between the codes is given in Appendix 7, and also on the relevant tables.

1. Obtaining mathematical information
    - A. The ability for formalized perception of mathematical material, for grasping the formal structure of a problem.

2. Processing mathematical information
    - A. The ability for logical thought in the sphere of quantitative and spatial relationships, number and letter symbols; the ability to think in mathematical symbols.

    - B. The ability for rapid and broad generalization of mathematical objects, relations, and operations.

    - C. The ability to curtail the process of mathematical reasoning and the system of corresponding operations; the ability to think in curtailed structures.

    - D. Flexibility of mental processes in mathematical activity.

    - E. Striving for clarity, simplicity, economy, and rationality of solutions.

    - F. The ability for rapid and free reconstruction of the direction of a mental process, switching from a direct to a reverse train of thought (reversibility of the mental process in mathematical reasoning).

3. Retaining mathematical information
    - A. Mathematical memory (generalized memory for mathematical relationships, type characteristics, schemes of arguments and proofs, methods of problem-solving, and principles of approach).

4. General synthetic component

A. Mathematical cast of mind.

Class 081

For the first half term the class observer played a passive role. By this we mean that he sat inconspicuously towards the back of the classroom with a seating plan and record sheet at hand so that every clue to ability that was available was recorded with the pupils name. Clues were classified as positive or negative according to whether the child showed a particular strength or weakness and were graded as High, Medium or Low. It was possible to categorise, in this way relative to class norms. For example we could establish where one pupils succeeded and another failed on similar work.

The teacher kept records of his/her own as well as he/she could. Lessons were usually in two parts, the first being class teaching where opportunities occured for pupils to show abilities in response to questions from the teacher or from their own questioning, and the second part being where the pupils worked on their own, attempting to solve mathematical problems based on the class discussion. During the second activity pupils would put up their hands to ask questions, usually when they were in difficulty. The teachers would give help as it was required. The observer continued the 'passive' role and supplemented what the teacher could do by answering some of these questions from individuals, recording any clues that were revealed. In this way we were able to record clues as they would naturally occur in the lessons.

Table 12.3 shows the levels that were confidently recorded by the observer during the first phase and Table 12.4 shows the records accumulated by the teacher over the same period (half a term) These tables can be compared with each other, and it is immediately apparent that an observer, with only one task in mind can record much more from the clues that come available than the teacher who is pre-occupied with all the other aspects of class activity.

TABLE 12.3

Checklist levels recorded by observer during 'passive' phase of observation (11 lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•			•	••		••													••		••					
08102						•					••											••					
08103				–	•	•			•		•						•	–		–							
08104				••			••			•	•							•		•					–		
08105							•				•											•					
08106	•			•									•														•
08107																					•	•					
08108	••			••	•	•	••		•	•	••		•			•	•		•	•	•				••		
08109						•	•				•									•							
08110						•																			•		
08111				•	•	•			•		–			•				•	•							•	–
08112						•					•									–		–			•		
08113	•			•						•	•			•													
08114						•					•									•							
08115						•					•			•				•				•					
08116						–				•										–	•	•			–		
08117	–			–			–				–							–			–	–	–				
08118						•				•	••												–				
08119				•	•						••							•		•							
08120	•			••	•		•			••	••	•	•	•				•	••	•	••	••	••				
08121				•		•	•				•							•				••					
08122	•			•	•	•														•							
08123	•			•	•		•		•	–																	
08124	•				•	•	•				•									•							
08125	•			•	•	•	•			•																	
	40	0	0	56	40	60	44	0	16	32	64	4	12	16	0	4	8	28	12	44	20	40	12	0	20	4	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.4

Checklist levels recorded by teacher during 'passive' phase (approximately half terms lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•			•					•	•	•											•					
08102																											
08103	•	•																•	•								
08104	•			•					•	•	•								•					•			
08105																											
08106																											
08107																											
08108		•				•	•		•	•								•	•					•			
08109																											
08110																											
08111				•										•	•												
08112				•																				•			
08113	•			•																							
08114																											
08115																											
08116																											
08117																											
08118																											
08119																								•			
08120				•							•		•	•	•									•			
08121	•	•		•					•	•	•		•							•				•			
08122																											
08123	•	•							•	•	•		•														
08124	•	•							•	•																	
08125																											
	28	20	0	28	0	4	4	0	24	24	20	0	12	8	8	0	0	8	12	4	0	4	0	24	0	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.5

Checklist levels recorded by observer after 'active' phase of observation (four further lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•			•	••		••		•	•	••	•	•	•				•		••		••					
08102	•			•	•	•			•			•		•	–			•		•		••					
08103	•			–	•	•	–		•		•						•	–		–							
08104	•			••			••			•	•							•		•					–		
08105	•				•		•		•		•		•	•			•		•	•		•					
08106	•			•	•				•	–	•		•	•			•		•								•
08107	•			•	•				•	–			–						•		•	•					
08108	••			••	••	••	••	•	•	•	••		•			•	••		•	•	••		••		•		
08109	•			•		•	••			•	•	•	•					•		•			•				
08110	•					•	••			•										•		•			•		
08111	•			•	•	•	•		•		•		•	•			•	•	•							•	–
08112	–			•	•	•	•				•			•					–	–		–			•		
08113	•			•	•					•	•			•						–							
08114	•			•	•	•			•	•	••		•		•	•		•		•				•			
08115	•				•	•	•			•	•			•				•		•		••					
08116	•			–	•	•	–			•	•	–	–	•	•		–	–	–	–	•	•		–	–		
08117	–			–	–		–				–	–	–					–		–	–	•	–	–			
08118	–			–	–	•				–	••							–	–	–			–	–			
08119	•			•	•		•				••	•	•		•			•	•	•						•	
08120	•	•	•	••	••		•			••	••	•	•	•				•	••	•	••	••	••	•			
08121	•			••		•	••				••		•				••	••	•	•	•	••		•			
08122	•			•	•	•					–		–						–	••							
08123	•			•	•		•		•	–	–						•	–	•						•		
08124	•				•	••	••				••		•	•				••	•	•				•			
08125	•			•	•	•	•			•	–		–		–	–		–					–	–			
	100	4	4	88	84	60	72	4	40	64	88	28	64	44	20	12	32	68	56	80	76	44	24	32	24	8	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.6

Additional clues recorded by teacher during 'active' phase (three weeks)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101							•		•												•						
08103	•													•							•						
08104				•	•		•						•						•								
08108	•			•			•						•								•						
08112							•				•	•															
08113	•			•	•	•																					

TABLE 12.7

Checklist levels recorded from work book scrutiny, accumulated over 'active' and 'passive' phases.

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6 a	6 b	6 c	7 a	7 b	8	9	10	11 a	11 b	12	13	14	15	16	17	18	19	20	21	22 a	22 b
08101	•					••	••			•										•		•					
08102	•					••	••			•										•							
08103	•					••	•			•																	
08104	•					••	•			•										•		•					
08105	•					••	•																				
08106	•					••	•			•										•		•					
08107	•					••	•			–										–							
08108	•					••	••													•		•					
08109	•					••	•			•			•							•							
08110	•					••	•			•										•							
08111	•					••	•			•										•							
08112	•					••	•			•										•							
08113	•					••	•			•										–							
08114	•					••	•			•										•					•		
08115	•					••	•			•										•		•					
08116	•					••	•			•																	
08117	•					••	–			•										•							
08118	•					••																					
08119	•					••	••			•										•							
08120	••					••	••						•							•		•					
08121	••					••	••			•										•		•					
08122	•					••	•			•										•							
08123	•					••	•			•										•							
08124	•					••	••			•										•		•					
08125	•					••	•																				
	100					[illegible]	96		[illegible]																		

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE

TABLE 12.8

Recategorisation of information from Table 12.3 to Table 12.7 for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2, 3)	2A (4,5, 6,21)	2B (7,8)	2C (9,10)	2D (11,12, 13,19)	2E (14, 19, 20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
[illegible]	•	••	••	•	•	•	••	••		10% 5%
[illegible]	•	••	•	•	•	•	•	••		10%
[illegible]	•	–	•		•	–	–	•		
[illegible]	•	••	•			•	•			
[illegible]	•	•	•	•	•	•	•	•		10%
[illegible]	•	•	•	•	•	•			[illegible]	10%
[illegible]	•	•	–	–		•	–	•		
[illegible]	••	••	••	•	••	•	••	••		5%
[illegible]	•	••	•	•		•	•			10%
[illegible]	•	••	•				•	•		
[illegible]	•	•	•	•	•	•	•		[illegible]	10%
[illegible]	–	•	•	•		–	–			
[illegible]	•	•	•		•		–			
[illegible]	•	•	••	•	•	•	•			10%
[illegible]	•	•	•		•	•	•	••		10%
[illegible]	•	–	•	–	–	–	–	•		
[illegible]	–	–	•	–	–	–	–	–		
[illegible]	–	–	•			–	–			
[illegible]	•	••	••	•	•	•	•		[illegible]	10%
[illegible]	••	••	••	•		••	•	••		5%
[illegible]	••	••	••	•	••	••	•	••		5%
[illegible]	•	•	–	–	•	–	•			
[illegible]	•	•	–		•	–	•			
[illegible]	•	••	••	•	•	••	•	•		5%
[illegible]	•	•	–	–	–	–				

Furthermore, the grading of observations that was possible for the observer was not possible for the teacher. The coding for tables based on the observer's records is ** = Grade H(high), * = Grade M(medium), - = Grade L (low). The single code for tables based on teachers records is: * = showed strength on this category'.

It is important to note that from the sample of classes observed, Class 081 produced the highest rate of clues. The teacher had a teaching style that involved a large proportion of the class in questions and answers as a topic was developed. Table 12.3 can be compared with later tables of a similar kind based on other classes. It is apparent, however, that it was not possible for the teacher himself to maintain adequate records.

The profiles gradually emerged for a number of pupils and it could be conjectured that given more time Table 12.6 would be closer to completion. There were some checklist items that were not observable at all during this phase and there were a considerable number of pupils on whom decisions could not be reached on a large number of items.

Table 12.3 provided a sound basis for the second phase of activity. This is where the observer played a more 'active' role. At appropriate times, when the pupils were working independently, the observer would spend a few minutes with pupils on whom certain categories of information has not been recorded. He would chat to them about the work and pose questions that they had to tackle on the spot, related to the levels of attainment on the current work. In this way further clues could be recorded. The observer was merely simulating what a teacher could do during a lesson. This active questioning by the teacher and positive feedback must be compared with the answering of a pupil's questions when they get into difficulty - a process resulting in negative feedback.

It was possible for the teacher to pursue this course of action to some extent at the same time, but for the purposes of the experiment the observer did most of the work.

As a result of the 'active' phase, Table 12.3 was updated to Table 12.5. Extra clues that the teacher recorded during this time appear in Table 12.6. Of course, at this stage, there were many more gaps in the teacher's own records than in the observers.

Table 12.5 evolved into a tool from which reliable identification could be made. Thus this teacher had a style which allowed adequate clues to be gathered from the classroom activity alone, even though he himself could not keep adequate records.

There were still some checklist items that could not be recorded, however. Items 2 and 3 - the ability to handle problems with redundant or insufficient information, item 6C - ability to solve problems of an unusual nature, item 12 - the ability to look for alternative methods and item 22 - the cast of mind, were all very scantily observed. This implies that, even in the 'active' phase, the type of the challenges set to the pupils did not allow conclusions to be reached on these items. In particiular, the lack of challenge coded as 6C is of concern when one considers the adequacy of high level challenges for children of high ability.

Table 12.7 records the checklists levels that were discovered from scrutiny of workbooks. The very narrow range of possibilities is quite apparent. Much of the work that pupils present in their workbooks is of a routine nature, either solving problems of a kind that the teacher has demonstrated in class (item 6a) or one on the same topic with just minor refinements of technique (item 6b). Not all of the pupils progressed to the 'hardert' questions, which could usually be categorised as 6b H; they could get by on the drill questions. When harder questions are attempted, it is not easy to associate the relevant checklist items with the pupils' work. First it would take a good deal of time to put the clues together for each child, and second, when the pupil is not present to discuss the work, it is difficult to deduce the route that led to what was finally written in the workbook. How long did he spend on it? Did he have help? Did he try different approaches?

It is important to consider, however, that most of the conclusions that teachers draw are based on records of attainment in workbooks. In this particular class we can tell from Tables 12.4 and 12.6 that it was not possible to draw valid conclusions from records of clues in the classroom. In contrast, the clues from the classroom, that the observer was able to record were both more reliable and more adequate than records from workbooks.

Table 12.8 shows the recategorisation into items based on the original source (the work of V.A. Krutetskii). We recall that various checklist items referred to similar Krutetskii categories. We should note that a broad spectrum of items were observable from pupils in this class, who were from the 13-14 year age group. Nevertheless some of the Krutetskii category levels could only be determined from a limited range of checklist items. For example, Krutetskii's Category 1, 'Gathering of Information', could only be determined from clues related to one of the three possible items on the checklist (1,2,3) to which it was related. We can conjecture that it was the kind of mathematical provision which dictated this limitation, so that though relative abilities in this group of children showed up - the identification of top 10% and 5% is given in Table 12.8 - the differences would be more marked if higher levels of challenge had been present.

School 03

The work in school 03 adds to our insights in a number of ways. First, we are dealing with a different kind of teaching environment. This is important in view of our discovery that differences of accuracy and bias of judgement do exist from teacher to teacher within a subject. Second, in school 03, we had the co-operation of teachers of both the top set and second set in mathematics. This made it possible to try out the checklist and assess its effectiveness as an aid to classification in two groups of pupils for whom there were different levels of expectation. This allowed us to assess the useability of the instrument in these two settings and also to consider the identification of the top 10% over the combined group. It may have been that some pupils of high ability were in the second ability set.

The conclusions are connected with passive observation of a sample of lessons over half a term on the part of the observer and on the accumulation of clues on the part of the teacher, over the same period, that emerged from the normal classroom routine.

We will report the results for the two classes separately.

Class 031 (The higher ability set)

Tables 12.9 to 12.12 contain the checklist levels that were recorded over this period. The style of teaching differed from the previous one reported (class 081). The class teaching and discussion was much more formal so that interaction between pupils and teacher were more limited. The class was certainly questioned during the discussion of the topic, but the bias was towards formal instruction rather than involved discussion when the class were actively encouraged to contribute to the development of the work. The rate of observation of clues during this phase of activity was thus lower for two reasons. First the class involvement was not so high and second because, this formal part of the lesson was shorter, being used as a device for moving the pupils onto exercises that they attempted individually.

During the second phase of the lesson the teacher answered individual problems that pupils had, but also found time to do some marking of work from which discussion could emerge.

The observer was able to do the same as the teacher during this second phase and it was from here that the majority of clues emerged. Table 12.9 can be compared with Table 12.3. The majority of clues for Table 12.9 came from the second phase of the lesson. The majority for Table 12.3 came from the first phase of the lesson.

The teaching styles produce similar amounts of clues, though the kinds in Table 12.3 associated with class discussion are more variable in nature than from Table 12.9 associated with individual discussion. Table 12.9 gives complete information for the class in certain categories. These are rather like the categories that can be observed when workbooks are scrutinised. These categories do, however, link with the kinds of questions that pupils are asked to attempt, limited in this class to drill and practice based on the methods demonstrated in the first part of a lesson.

TABLE 12.9

Checklist levels recorded by observer from class observation ('Passive' ten lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03101	•			•		•	•		•	•	•											••			•		
03102	•			•	•	•	••		•	•	•									•					•		
03103	–					•	–		•											•					•		
03104	•					•	–		•	•	•									•					••		
03105	•					••	•		•	•	•			•						•					•		
03106	–					•	–		•											•					•		
03107	••			•	••	••	••		••	•	••							•	•	•		••			••		
03108	•					•	–		•											•							
03109	•					•	••		•	•	••									•					•		
03110	•			•		•	•		•	•								•		•					•		
03111	••					•	•		••	•										••					•		
03112	•			•		•	–		•	•	–							•		•		•			•		
03113	•			•		•	–		•	•										•		•			•		
03114	•			•		••	•		•	•	•									•					•		
03115	••	•		•	••	••	••		••	•	••		•				•		••	••		•			••		
03116	–					–	–		•	•	•					–			–	–					•		
03117	•			•		–	–		•									•		•		•			•		
03118	•			•		•	•			•	••									•							
03119	•			•	•	•	•		•	•	••		•							•					•		
03120	•			•		••	•		•	•	••									•					•		
03121	•					•	•		•											•		•			•		
03122	•			•		••	•		•	•	••									••					•		
03123	•					•	•		–											–							
03124	•					•	•		•		•									•					•		
03125	•					•	•		•											•					•		
	100	4	0	52	16	100	100	0	96	68	60	0	8	0	0	4	4	16	12	100	0	28	0	0	88	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.10

Checklist levels recorded by teacher (Approximately half term)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03101									●										●		●						
03102		●			●	●	●			●	●	●		●	●				●		●		●		●		●
03103																											
03104						●			●		●		●						●		●		●	●	●		
03105																					●		●	●	●		
03106																			●						●		
03107		●			●	●	●		●	●	●	●	●						●	●	●		●	●	●		●
03108																											
03109									●												●				●		
03110				●									●								●				●		
03111		●								●									●		●			●	●		
03112																									●		
03113				●									●												●		
03114																			●						●		
03115		●			●	●	●			●	●	●	●						●	●	●		●		●		
03116											●								●						●		
03117																											
03118									●		●								●						●		
03119		●							●												●				●		
03120																			●						●		
03121									●																●		
03122		●			●					●	●								●		●				●		
03123																											
03124						●	●		●	●						●			●		●		●		●		
03125																									●		
	0	24	0	8	16	20	20	0	32	24	28	12	20	4	4	4	0	0	52	8	48	0	24	16	80	0	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.11

Checklist levels recorded from workbook scrutiny checklist codes

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6 a	6 b	6 c	7 a	7 b	8	9	10	11 a	11 b	12	13	14	15	16	17	18	19	20	21	22 a	22 b
03101						•				•																	
03102	••			•		••	••				•									•							
03103						•																					
03104						••	•													•							
03105						••	•													•							
03106						•	–													–							
03107	••					••	••			•										•							
03108						•	–																				
03109	•					••	•			•										•							
03110						••	•																				
03111	•					••	••													•							
03112	•					••																					
03113						•	•																				
03114						••	•																				
03115	••					••	••			•										•							
03116						•														–							
03117						•																					
03118	•					••	•			•										•							
03119	•					••	•						•														
03120	•					••	•													•							
03121	•					••	•													•							
03122	•					••	••													•							
03123						•	–													–							
03124						••	••													•							
03125	•					••	•													•							
	48	0	0	4	0	100	80	0	0	20	4	0	4	0	0	0	0	0	0	64	0	0	0	0	0	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.12

Recategorisation of information from Table 12.9 to Table 12.11 for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2,3)	2A (4,5,6,21)	2B (7,8)	2C (9,10)	2D (11,12,13,19)	2E (14,19,20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
03101	•	•	•			•		••		Top 10%
03102	•	••	•	•	•	•	•	•	ANAL/LOG	Top 5%
03103	–	–	•			•	•			
03104	•	–	•	•		•	•	•		
03105	•	•	•		•		•	•		Top 10%
03106	–	–	•			•	–			
03107	••	••	••	•	•	•	•	••	ANAL/LOG	Top 5%
03108	•	–	•				•			
03109	•	•	•			•	•			
03110	•	•	•	•		•	•			
03111	••	••	••			•	••			Top 5%
03112	•	–	–			•	•	•		
03113	•	–	•	•			•	•		
03114	•	•	•				•			
03115	••	••	••	•	•	••	••	•		Top 5%
03116	–	–	•		–		–	–		
03117	•	–	•			•	•	•		
03118	•	•	••				•			Top 10%
03119	•	•	••	•			•	•		Top 10%
03120	•	•	••			•	•			Top 10%
03121	•	•	•				•	•		
03122	•	••	••			•	••	•		Top 5%
03123	•	–	–				–			
03124	•	••	•		•	•	•	•		Top 10%
03125	•	•	•				•			

Categories 6a and 6b are associated with these routines. Category 1 is associated with the ability to understand information, but the abundance of single *'s in this column indicates that high level challenges were not in general presented. Item 21 concerns aptitude in arithmetic - again not challenged at a high level. One might suppose that the clues could be much stronger and variable either if the challenges were more adequate and variable in the work set to the class as a whole or in the discussion between individual pupil and teacher.

It is interesting that the amount of information that the teacher himself recorded was much fuller than in class 081 and this is a consequence of the time devoted to work with individuals. There is a little more variety in Table 12.10 than Table 12.9, but there are signs in both tables of more complete knowledge of some pupils than others.

The same conclusions about inadequacy of the clues that can be used from workbook scrutiny is apparent from Table 12.11.

If one were to attempt a classification of pupils into top 10% and next 10% (in this case top and bottom halves of the class) the lack of the full range of clues, particularly from the teacher's own records would make the exercise tentative, yet when the three tables are combined and reclassified into Table 12.12, we see that from this 'passive' activity a preliminary classification was possible. If we had then pursued an 'active' phase where we worked closely with pupils on whom knowledge was rather scant, we could improve the table of results and carry out a more confident classification.

It is important to repeat that the range of clues that combined to give the categorisations of Table 12.12 was limited by the level of provision. There must be a source for concern if the sort of work presented to pupils does not challenge the full range of aptitudes, both because of mathematical growth of the children and of their identification. We must remember too that the tentative identification that was possible from Table 12.12 relied heavily on the categories discovered by the observer. Thus there were enough clues possible for an identification but the teacher had been too preoccupied to record a sufficient number.

Class 032 (Second ability set)

The style of teaching of this group was much the same as for 031. The marked difference was in the level of challenge set to the pupils. The teacher was at pains to make sure that all pupils could gain a working knowledge of the topics covered[1]. This inevitably led to an algorithmic approach and to routine exercises. This was understandable in that many of the class achieved success on nothing but routines.

The formal part of the lesson did allow for some variety of clues to be picked up because the teacher looked to the class for ideas and involved a large number of pupils.

The classification of clues was relative to the levels of challenge set within this group, so that norms were different from that for class 031. For this reason it was difficult to establish cross group norms, though it did appear that the best of class 032 would not be a candidate for the top 10%.

The exercise was to try to divide the class into two equal groups on mathematical ability. In this way we could establish whether the checklist was an effective instrument in this lower ability environment.

The clues that could be observed were inevitably of the kind associated with the routine work that the children were set and these routines made it impossible to get very much from workbooks. The dominant item here was 6a, associated with solving problems by means of an algorithm. There were very few clues to any other aspect of mathematical ability, including the ability to understand information. By way of contrast, Item 1 was usually observable from the workbooks of higher sets.

Again the teacher recorded an inadequate amount for proper identification, but added to the observer's records and when all clues were reformulated into Table 12.16 we see that a spectrum of abilities could be built up for a number of pupils.

TABLE 12.13

Checklist levels recorded by observer from classwork
(Passive role, Ten lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6 a	6 b	6 c	7 a	7 b	8	9	10	11 a	11 b	12	13	14	15	16	17	18	19	20	21	22 a	22 b
03201	*			**	**	**	**			*	**					*				**					*		
03202	*			*	**	*					**									**					*		
03203	*					-				-	-														*		
03204	*					*	-		*	-									-					-			
03205	*					*	*																		*		
03206	*			*	-	*	-				-									-					*		
03207	*				*	**	*			*	**									*					*		
03208						*	*		*	-										*					*		
03209	*			*		-														-					-		
03210	*			**	**	**	*				*								*	*					*		
03211	*			*	-	-	-				-									-					-		
03212						*	*																		*		
03213	**			**	**	**	**						**			**	**			**					**		
03214	*			*	-	-	-				-									-					-		
03215	-			*	*	-																			*		
03216	*				*	*	*				**									**					*		
03217		*			-	-					-									-					*		
03218				*	*																						
03219	*			**	-	-	-				-									-					-		
03220	*				*	**	**				**									*					*		
03221				*		-					-		*			*	*					-					
03222	*			*	-	*	*				-														*		
03223				*		*	*	-												-							
03224	-				-	-	-				-												-				
03225	*			*	*	*	*			-								*		*					-		*
03226	-				-	-	-				-						-			-					-		
03227	*				*	*	*													*					*		
	78	4	0	56	70	96	70	4	7	19	59	0	7	0	0	7	11	4	7	67	0	4	4	4	81	0	

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.14

Checklist levels recorded by teacher
(Approximately half term)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03201	*	*		*	*				*	*							*		*				*	*			*
03202	*	*		*									*						*					*			
03203																			*								
03204																											
03205																			*								
03206				*															*					*			
03207																	*		*				*				
03208			*						*										*				*	*		*	
03209																											
03210																											
03211																											
03212	*		*	*	*				*	*			*				*		*					*			*
03213	*	*		*					*				*				*		*				*	*			
03214																			*								
03215																											
03216																								*			
03217																	*							*			
03218				*	*														*				*	*			*
03219		*	*	*																				*			
03220																	*							*	*		
03221																								*	*		
03222			*	*														*	*					*		*	
03223																											
03224																											
03225		*	*	*	*				*		*	*	*			*			*				*	*	*		*
03226																			*					*	*		
03227				*	*				*			*	*				*		*					*	*		
	[illegible]	19	19	37	19	0	0	0	[illegible]	7	[illegible]	[illegible]	19	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	22	[illegible]	19	7	15

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.15

Checklist levels recorded from workbook scrutiny

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03201						* *	*													*					*		
03202	*			* *		* *	*													*					*		
03203				-		*	-																				
03204						*	-																				
03205				-		*	*																				
03206						*																			*		
03207		* *	* *	*		* *	*			*																	
03208				*		*														*							
03209				-		-	-																				
03210				*		*	*													*							
03211				-		*																					
03212						*																			*		
03213						* *	*																		*		
03214				-		*																					
03215				*		*														*							
03216				*		*	*																				
03217				-		-																					
03218				*		*																					
03219				* *		*																					
03220	* *	* *	* *	* *		* *	* *													*							
03221				-		-																					
03222				*		* *																					
03223		* *	* *	*		* *																					
03224				-		-	-																				
03225	*			*		* *	* *																		*		
03226				-		-	-																				
03227		* *	* *	*		*	* *																				
	11	15	15	48	0	100	37	0	0	4	0	0	0	0	0	0	0	0	0	22	0	0	0	0	33	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.16

Recategorisation of information from Table 12.15 to Table 12.15
for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2,3)	2A (4,5,6,21)	2B (7,8)	2C (9,10)	2D (11,12,13,19)	2E (14,19,20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
03201	*	* *	* *		*	*	* *			TOP 1/5 of class
03202	*	* *	* *	*		*	* *		Anal/Log	TOP 1/4 of class
03203	*	–	–			*				
03204		–	–			–	–			
03205	*	*				*				
03206	*	*	–			*	–			
03207	*	* *	* *		*	*	*			TOP ½ of class
03208	*	*	–			*	*		Spat/Geom	
03209	*	–	–				–			
03210	*	* *	*			*	*			TOP ½ of class
03211	*	–	–				–			
03212	*	*	*	*	*	*				TOP ¼ of class
03213	* *	* *	*	* *	* *	*	* *			TOP 12% of class
03214	*	–	–			*	–			
03215	–	–					*			
03216	*	*	*				* *			
03217	*	–	–		*		–			
03218		*			*	*			Anal/Log	
03219	*	–	–				–			
03220	* *	* *	* *		*		*			TOP 1/4 of class
03221		–	–	*	*			–		
03222	*	*	–			*			Spat/Geom	
03223	* *	*					–			
03224	–	–	–							
03225	*	*	*	*	*	*	*		Anal/Log	TOP ⅓ of class
03226	–	–	–			*				
03227	* *	* *		*	*	*	*			TOP ⅓ of class
									U	

There is enough evidence to suggest that identification by this means was possible if individual pupils had been challenged more adequately with appropriate material. Those who appeared to be in the top quarter of the class could then be compared in with the higher class on an individual basis.

The same important concern is evident, however, that the teacher's records were far less complete than the observer's.

School 07

This school had two 'parallel' top sets in mathematics so that the two groups were observed separately and a final identification of the top 10% (approximately one quarter of each set) was then done in consultation with both teachers.

The school adopted the policy of diverting pupils into a first and second ability set after their third year so that an important policy was for the two third year sets to be taught the same topics in much the same way. The classes took many of their exercises from work sheets which could be used by either class and were also periodically tested for progress, the same test material being used for both classes.

The teaching style used in both classes was to lead into the worksheets through class teaching. The sheets themselves contained questions that were designed to take the pupil progressively through the topic. The classes were involved to a limited extent with the formal part of the lesson as the teacher threw out questions to the class or to an individual. The teacher was available to sort out difficulties when the pupils worked as individuals for the latter part of the lessons.

For both these classes, 'passive' observation was used for the first half of term, followed by 'active' observation for a sample of lessons in the second half of term. The contrast between the clues to inability that could be discovered in the passive role, when the pupil/teacher interaction was governed by a pupil's difficulties, must be compared with the clues to ability that are discovered in the active role, when a pupil is presented with challenges appropriate to a particular mathematics aptitude in relation to the topic that is being developed in the classroom, and insights into his/her strengths is possible. Indeed the teacher would tend to obtain clues from particular kinds of pupil: those outgoing in character and/or those in difficulty. It must be stressed that in these classes as elsewhere this activity does not occur naturally and must be considered to be the key to successful identification with checklists.

Class 071

Tables 12.17 to 12.21 summarise the checklist levels that were recorded for this class. The extreme shortage of information contained in Table 12.19 is quite apparent. The task of record keeping and completion of checklists has been fully understood, as it had with all teachers and the teacher was keen to do the work. We can only conclude that this is a striking example of how difficult the task is for a teacher who is preoccupied with other classroom tasks both in noticing relevant clues and memorising them.

Table 12.17 shows what the observer recorded over a sample of lessons during the 'passive' phase. Though these are far too inadequate for the identification to be done, there was far more available than the teacher could record.

Table 12.18 was developed from Table 12.17 by the observer looking where gaps of knowledge existed and finding ways of filling those gaps by questioning pupils at their work.

Table 12.20 illustrates the narrow range of clues that could be confidently recorded from workbook scrutiny, though within this narrow range confirmation could be found for assessments made in the classroom.

The accumulation of evidence into Table 12.21 led to confident identifications of the top 10% and 5%.

TABLE 12.17

Checklist levels recorded by observer from classroom observation during 'passive' phase (eleven lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07101																											
07102											•														•		
07103																									–		
07104										•	•														•		
07105						•	•			–	•	–													–		
07106	•			•	••	••	••		•		••									•			••		••		
07107	•			•	••																						
07108																						••					
07109	••				••	••	••									•											
07110	••					••	••									•	•		•	•							
07111	•																	•	•	•							
07112	•									–	–							–		•					–		
07113						•	–			–	–														–		
07114						–																					
07115	••			•		••				•	••		•						•		•	•			–		
07116											–																
07117	•					••										•		•									
07118	•				•								•	•				–									
07119	•				••	••																			•		
07120	••				••	••	••				••									•					••		
07121	••					••				•	••									••							
07122	–									–	–								–			–					
07123																											
07124																											
07125					–	•																–					
07126	–																										
	88	0	0	12	85	96	84	0	4	23	58	8	33	19	28	28	8	23	15	48	46	65	52	19	38	19	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.18

Table 11.8 updated to include checklist levels recorded by observer during 'active' phase (nine lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07101	-				-	*	-						-								-	-			-		
07102	*				*	*	*				*		*		*			*		*	*	*			*		
07103	*				-	*	*				-		-		-			-		-	-	-			-		
07104	*				*	**	*			*	*	-	-	*		-	-	-		-	*	*			*		
07105	-				*	*	*			-	*	-	-	*		-	-	-	-	-	-	-	-		-		
07106	**			*	**	**	**		*		**	*	*	*		*	*		*	*	*	**	**		**		
07107	*			*	**	**	*			*	*				*			*		*	*	*	*		*		
07108	*				**	**	**				*		*								*	**	**		**		
07109	**				**	**	**				*		*	**		*	*	*		*	*	*	**	*	**		
07110	*				**	**					*		*			*	*	*	*	*	*	*		*	**		
07111	*				**	**					*						*	*	*	**		*	*		*		
07112	*				*	*	-			-	-							-	-	*	-	-	*		-		
07113	*				*	*	-		-	-			-			-	-	-			-	-	*		-		
07114	-				-	-				-								-		-	-	-			-		
07115	**			*	**	**	*			*	**		*		*		*	*	*	**	*	*	*	*	-		
07116					-	*					-									-	-	-			-		
07117	*				*	**	**				*					*		*		*	*	*			*		
07118	*				*	*					*		*	*				-		*	*	*			*		
07119	**				**	**	*				*					*			-		**	**	*		*		
07120	**				**	**	**				**		*		*	*	*			*	*	*	*		**		
07121	**				**	**	**			*	**		*		**		*			**	*	*	**	*	**		
07122	-			-	-		*	-		-	-							-	-	-	-	-			-		
07123						*															-	-		-			
07124	-				*	*	-				-							-		-	-	-			-		
07125	-				-	*					-							-		-	-	-			-		
07126	-				-	*	-													-	-	-			-		
	92	8	0	15	96	96	27	4	8	35	81	12	54	19	23	31	58	50	27	81	96	100	96	15	96	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.19

Checklist levels recorded by teacher (during the term of observation).

CHECKLIST CODES

Pupil codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07101																											
07102																											
07103																											
07104																											
07105														•													
07106																							•		•		
07107																											
07108																							•				
07109																						•	•				
07110																											
07111																						•	•		•		
07112																											
07113																											
07114																											
07115																											
07116																											
07117																											
07118																											
07119																											
07120																						•	•				
07121																					•		•				
07122																											
07123																											
07124																											
07125																											
07126																											

TABLE 12.20

Checklist levels recorded from workbook scrutiny over the complete term

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07101	–					•	–														–	–			–		
07102	•					••	•						•				•				•	•			–		
07103	•					•	•														–	–			–		
07104	–					••	–				•										–	–			•		
07105	–					•	–														–	–			•		
07106	•					••	•														•	••			••		
07107	•					••	•														•	•			•		
07108	•					••	•						•								•	••			••		
07109	•					••	•														•	•			•		
07110	••					••	••				•										•	••			•		
07111	••					••	••				•										•	••			•		
07112	•					•	–														–	–			–		
07113	•					•	–														–	–			–		
07114	–					•	–														–	–			–		
07115	•					•	•														•	•			–		
07116	–					•	–														–	–			–		
07117	•					••	•														•	•			•		
07118	•					••	•														•	•			•		
07119	•					••	•														•	•			•		
07120	•					••	••				•		•								••	••			••		
07121	••					••	••														••	••			••		
07122	–					•	–														–	–			–		
07123	•					•	–				•										–	–			•		
07124	–					•	–														–	–			–		
07125	–					•	–														–	–			–		
07126	–					••	–														•	–			•		
	100	0	0	0	0	100	100	0	0	0	19	0	12	0	0	0	4	0	0	0	100	100	0	0	100	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.21

Recategorisation of information from Table 12.17 to Table 12.20 for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2, 3)	2A (4,5, 6,21)	2B (7,8)	2C (9,10)	2D (11,12, 13,19)	2E (14, 19, 20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
07101	-	-		-				-		
07102	*	*	*	*	*	*	*	*		
07103	*	-	-	-	-	-	-	-		
07104	-	*	*	-	-	-	-	*		
07105	-	*	-	-	-	-	-	-		
07106	**	**	**	*	*	*	*	**		Top 5%
07107	*	**	*		*	*	*	*		Top 10%
07108	*	**	*	*		*	*	**		Top 10%
07109	**	**	*	*	**	*	*	*		Top 10%
07110	**	**	*	*	*	*	*	*		Top 10%
07111	**	**	*			*	**	**		Top 10%
07112	*	-	-			-	*	-		
07113	*	-	-	-	-	-		-		
07114	-	-	-			-	-	-		
07115	**	**	**	*	*	*	**	*		Top 5%
07116	-	-	-				-	-		
07117	*	**	*		*	*	*	*		Top 10%
07118	*	*	*	*	*	-	*	*		
07119	**	**	*		*	*		**		Top 10%
07120	**	**	**	*	*	*	*	*		Top 10%
07121	**	**	**	*	**	*	**	*		Top 5%
07122	-	-	-			-	-	-		
07123	*	-	*			-		-		
07124	-	-	**			-	-	-		
07125	-	-	-			-	-	-		
07126	-	-					-	-		

Class 072

Tables 12.22 to 12.26 contain the comparable information for class 072. We see that the teacher was again only able to give confident conclusions for a limited number of pupils and that inadequate evidence was available from the 'passive' classroom observation. Again, however, the 'active' phase enabled sufficient evidence to be accumulated to enable confident identifications to be carried out.

We note that in both classes 072 and 073 the sorts of challenges normally present from which clues to ability could emerge were not such that the full range of aptitudes could be observed. Flexibility of the mental processes was not being challenged, in particular.

We can contrast the clues that were available for these classes in the normal way with those available in class 081 where variety of challenge was evident. It could be a consequence of having two rather broad ability sets in school 07 that this is so and a price that is paid for keeping all of these children at a common level of experience, prior to making decisions on final setting in their fourth form. It was possible, because of the similarity of work for classes 071 and 072, to make comparisons between pupils in the two sets, unlike the two classes of school 03. It was also possible to draw more extreme comparisons between the higher and lower ability pupils than was possible in classes with a more limited range of abilities.

Advantages of this sort must be set against the disadvantages of a lower level of challenge for the most able that was available from examination orientated routine work.

TABLE 1[illegible]

Checklist levels recorded by observer during 'passive' phase
(Ten lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07201	* *					* *	* *											*									
07202						*	-											-									
07203																											
07204	*					*	*											-									
07205	* *			* *												*	*	*				*					
07206						*																					
07207	-						-					-						-									
07208																											
07209	*					* *	-											-									
07210																					-	-					
07211	* *					* *	*											*									
07212	* *					* *												*				*					
07213						* *	* *				*									*		*			*		
07214	* *						* *										*			*	* *	* *			* *		
07215	*					*	*											*									
07216																											
07217	*			-		*					-								*	-							
07218	*					*	*											*	-	-					-		
07219											-																
07220	*			*	*	*	-				*							*									
07221	*										-														-		
07222																											
07223																											
07224	*					* *	* *											*		*					*		
07225																			*		*	*					
07226	-					-	-			-								-							-		
07227	* *						* *											* *		*		* *			*		
07228	*			*	*	* *	*				-							*									
07229	*						*				-					-	-	-	*	-		-			*		
07230	*						*											-									
	33	0	0	13	7	50	57	0	0	3	27	0	0	0	0	7	10	37	13	23	10	27	0	0	27	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.[illegible]

Checklist levels which include those recorded during 'active' phase (Further Seven lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07201	* *				*	* *	* *				*							*	*	*		*			*		
07202	*	–				*	–				–							–	–	–					–		
07203	* *	* *	*			* *					*								*	*					*		
07204	*	–				*	*				*							–	*	*					–		
07205	* *			* *		*					*					*	*	*	*	*		*			*		
07206	*					*	–				–								–	–					–		
07207	–	–		–		* *	–				–		–	*	–			–		*			–		–		
07208	–	–		–		*	–				–		–	*	–					–	*	*	–		*		
07209	*	*				* *	–				*		–					–		*					–		
07210	–	–		–		*	*			–	–		–	*	–			–		–	–	–			–		
07211	* *	* *		* *		* *	*				*		*					*		*	*	*					
07212	* *	* *		* *		* *			*		*		–	*				*		*	*	*			*		
07213	*			*		* *	* *			*	*		*		*					*	*	*		*	*		
07214	* *	* *		* *		* *	* *			*	*		* *		* *		*	*		*	* *	* *		*	* *		
07215	*					*	*											–			*	*			–		
07216	*	–				*	*														*	*			–		
07217	*	–		–		*	*				–		–					–	*	–					–		
07218	*	–		–		*	*				–		–					*	–	–					–		
07219	*					*	*				–								–	–					–		
07220	*	–		*	*	*	–				*							*			*	*					
07221	*				*	*	–				–								–	–	*	*			–		
07222	*				*	*	*				*								*	*	*	*			*		
07223	*				*	*	*		*	–	*	–	–							*					–		
07224	* *				* *	* *	* *		*	*	* *		* *					*	*	*					*	*	
07225	*	*		*	*	*	*				*			*	*	*			*	*	*	*			*		
07226	–	–			–	–	–			–	–		–					–							–		
07227	* *	* *		* *	* *	* *	* *		*	*	* *	*	*	*		*	* *	* *		*	* *	* *			*		
07228	*	*		*	*	* *	*		*		* *	*	*					*		*	*	*			*		
07229	*	*		*		* *				–	–		–			–	–	–	*	–		–			*		
07230	*	–		*		* *	*				–		–					–							*		
	100	63	3	47	33	100	90	0	17	27	90	10	50	20	20	13	17	63	57	77	57	43	7	7	100	3	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.2b

Checklist levels recorded by teacher during whole period of observation (One Term)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07201																							*		*		
07202																											
07203	*																										
07204																											
07205	*			*																				*			
07206	*			*																							
07207	*																										
07208																											
07209																							*	*			
07210	–																							*			
07211																											
07212	*	*	*	*	*	*															*	*		*	*		
07213	*			*	*																	*		*			
07214	*	*	*	*	*	*			*										*		*	*	*	*	*		
07215	*																						*				
07216	*	*		*	*				*												*	*	*				
07217																											
07218																											
07219																							*				
07220	*																				*		*	*			
07221	*			*	*																*						
07222	*																				*						
07223																											
07224																											
07225																							*				
07226																											
07227	*		*	*	*				*		*								*		*	*	*		*		
07228	*	*	*	*	*	*			*		*	*	*						*		*	*	*	*	*		
07229	*			*																				*	*		
07230	*			*																				*	*		
	37	17	13	37	23	10	0	0	13	0	7	3	3	0	0	0	0	0	10	0	27	20	33	33	23		

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.25

Checklist levels recorded from workbook scrutiny
(One Term)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
07201	**					**	*														*	*			*		
07202	*					*	_				_										_	_			_		
07203	*	*	*			*	*				*		*												*		
07204	*					*	_														_	_					
07205	*					**	*																				
07206	**					*	_														_	_					
07207	_					**	_																				
07208	*					**	*														*	*					
07209	*					**															_	_					
07210	_					**	_															_					
07211	*					*																					
07212	**					**	**												*		*	*					
07213	*					*	*				*										*	*					
07214	**					**	**				*										*	*					
07215	*					*	_														*	*					
07216	*					**	*														*	*					
07217	*					*															_	_					
07218	*					**															_	_					
07219	*					*	*																				
07220	*					*					*										*	*					
07221	*					**					_										*	*					
07222	*					*																					
07223	*					*	*																				
07224	**					**	*																				
07225	*					*	*																				
07226	_					*	_														_	_					
07227	**					**	**				*								*		*	*			*		
07228	**					**	**				*										*	*			*		
07229	**					**	*																		*		
07230	*					**	*																				
	100	3	3	0	0	100	77	0	0	0	27	0	3	0	0	0	0	0	7	0	60	53	0	0	10	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12,26

Recategorisation of information from Table 12,22 to Table 12,25

for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2, 3)	2A (4,5, 6,21)	2B (7,8)	2C (9,10)	2D (11,12, 13,19)	2E (14, 19, 20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
07201	* *	* *	*			*	*	*		TOP 10%
07202	*	-	-			-	-	-		
07203	* *	*	*		-	*	*			
07204	-	-	*			-	*	-		
07205	* *	*	*		*	*	*	*		TOP 10%
07206	*	-				-	-	-		
07207	-	-	-	-	-	-	*			
07208	-	*	-				-	*		
07209	*	*	*	-		-	*			
07210	-	-	-	-	-	-	-			
07211	* *	* *	*	*		*	*	*		TOP 10%
07212	* *	* *	*		*	*	*	*		TOP 10%
07213	*	* *	*	*	*			*		TOP 10%
07214	* *	* *	*	* *	* *	*		* *		TOP 5%
07215	*	-				-	*			
07216	-	*	*					*		
07217	-	-	-	-		-	-			
07218	-	-	-	-		-	-			
07219	*	*	-				-			
07220	-	-	*			*	*	*		
07221	*	-	-			-	-	*		
07222	*	*	*			*	*	*		
07223	*	*		-			*			
07224	* *	* *	* *	* *		*	*			TOP 5%
07225	*	*	*		*	*	*	*		
07226	-	-	-	-		-				
07227	* *	* *	* *	*	* *	* *	*	* *		TOP 5%
07228	*	* *	* *	*		*	*	*		TOP 10%
07229	* *	*	-		-	-	-	-		
07230	-	*		-		-				

Class 010

This was a top set class and the top 10% was to be identified as the top half of this set. The teacher was not a mathematics specialist and was very concerned about lack of experience with able mathematics pupils. She was, however, very enthusiastic to be involved in the work and to use the checklist as an instrument in the identification process.

We can see from Table 12.29, however, that the constraints of inexperience coupled with the constraints of classroom management prevented adequate clues from being recorded even though Tables 12.27 and then 12.28 from the 'passive' and 'active' phases of observation show that adequate records were possible that would enable Table 12.31 to be compiled prior to confident identification of both the top 10% and top 5%.

The 'active' phase incorporated a further experiment with this particular class. A sample of lessons were designed with the checklist in mind so that the challenges and questions set to the class could incorporate the full checklist range. The variety of clues that emerged can be compared with the variety from class 081 reported earlier. It was possible to keep records of responses by individual pupils because of the way the lessons were prepared (unlike the insufficient records kept for class 081). Clues gained from this activity are included in Table 12.28.

Again, the confident clues that could be gained from workbook scrutiny were of a very limited kind.

TABLE 12.27

Checklist levels recorded after 'passive' phase of observation (fourteen lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
10101							•				•	–											•				
10102	–			•			–		•													•					
10103				–		•					–																
10104		–				–					–							–	–			–					
10105				•	•				•	–	•																
10106	••			•			•			•	•														•		
10107	••			•	•		••	•		••	•																
10108	•					•					–																
10109					•						–						–	•									
10110	••				••		•			•	•		•				•	••	••							•	
10111				•		•													•								
10112				•																							
10113	•			•	•	•	–			–	–															–	
10114							–																				
10115	–				–		–			–	–																
10116	•			•	•	•	–		•	–	–						–	–			–	–			–		
10117				•	–	•				–	–						–						•				
10118																											
10119	–	–		–	–	•	–			–											•						
10120					•		•		•																		
10121																											
10122	–					•	–			–	–																
10123	•			•		•	•				•																
10124						••	•				•																
10125	•				•	•	•		•	–	•																
10126																			•								
10127					•	•				•																•	
10128						•	•								•											•	
10129	••			••			••	••			•							••	•		••	•					
10130	•																										
10131	–			–	–	–	–		•	•	–		–				–	–		–					–	•	
10132	–				•	–	–				–	–				–	–	–					–				
10133					•	•			•																		
10134	•			•			•				•																
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.28

Checklist levels accumulated after 'active' phase of observation (a further nine lessons).

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
10101	·			·	·	··	·			·	·							·				·			·		
10102	–			·		·	–		·	–			–		–			·				·			–		
10103	·			–		·	–			–	–	–	–					·					·		·		
10104	·	–		–		·	–			–	–							–	–						–		
10105		L	E	F	T		C	L	A	S	S																
10106	··			·	·	··	··		·	·	·		·	·	·			·					··	·	·		
10107	··			·	·	··	··		··	··	··		·					·					··	·	·		
10108	·			–	–	·	–			–	–				–			–					–	–	–		
10109	·		–	·	·	·					–						–	·						·	–		
10110	··			·	··	··				··	··		·			·	·	··	··		·	·	·	··			
10111	·			··	··	··	··			··	··		·					·	·			·				·	·
10112	·			·	·	··	··			··	··	·	·	·			·	·				·	·				
10113	·			·	·	·	–		–	–	–							·		–	·				·		
10114	·			·		·	–			–								–							–		
10115	–			–	–	·	–			–	–							–									
10116				·	·	·	–		·	–	–						–	–			–	–		–	–		
10117	·			·	–	·	·			–	–							–		·	·		·				
10118	·			–		·	·			–	–							–							–		
10119	–			–	–	·	–			–	·							·		–	·				–		
10120	·			·	·	··	··		·	·	··			·				·			·				·		
10121	·			··	·	··	··		·	·	·		·		·			··			·		·	·			
10122	–			–		·	–			–	–					–		·	–						–		
10123	··			··	··	··	··		··	·	··			·				·		·			··		··		
10124	·			·		··	··				·							·		·	·						
10125	·			–	·	·	·		·	–	·							·				–				–	
10126	·			·	·	··	··		··	·	·							·	·	·	··	··	·	·			
10127	·			·	··	··	··			·	··		··					·		·						·	
10128	·			·	·	··	··			··	·				·			·					·	·		·	
10129	··			··	··	··	··	··		·	·		··					··	·	·	··			··	·	·	
10130	·			·		·	·				·							–									
10131	–			–	–	–	–		·	·			–				–	–		–					–	·	
10132	–			–	·	·	–				–	–				–	–	–					–				
10133	·			·	·	·	·		·	·	·			·				·			·				·		
	100	9	0	100	75	77	100	9	48	88	96	9	36	16	16	9	19	100	10	28	16	25	38	17	59	22	5

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 2.29

Checklist levels recorded by teacher (school 101) over one term.

CHECKLIST CODES

Pupil Code#	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
10101																											
10102																											
10103																											
10104																											
10105																											
10106																											
10107																											
10108																											
10109																											
10110																											
10111																											
10112																											
10113																											
10114																											
10115																											
10116																											
10117																											
10118																											
10119																											
10120																											
10121																											
10122																											
10123																											
10124																											
10125																											
10126																											
10127																											
10128																											
10129																											
10130																											
10131																											
10132																											
10133																											
10134																											

TABLE 12.30

Checklist levels recorded from workbook scrutiny (over one term).

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
10101	•					••	•				•							•									
10102	–					•	–				–							•									
10103	•					•	–											•									
10104	•					•	–			–									–								
10106	••					••	•											•									
10107	••					••	••				•							•									
10108	•					•																					
10109	•					•	•				•							•									
10110	••					••	•				•							•			•						
10111	••					••	•				•							•									
10112	•					••	•			•	•							•									
10113	•					•															•						
10114	•					•	–																				
10115	–					•	–																				
10116	•					•	•																				
10117	•					•	•				•							•									
10118	•					•	•																				
10119						•	–				–										–						
10120	•					••	•				•		•					•			•						
10121	•					••	•			–	•							•									
10122						•												•									
10123	••					••	•				•							•									
10124	•					••	•				•							•									
10125	•					••	••																				
10126	•					••	•				•							•									
10127	••					••	••				••		••					•									
10128	••					••	••				•							•									
10129	••					••	••				••		••					•									
10130	•					••	•				•																
10131						•																					
10132						•	–																				
10133	•					••	•																				
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.31

Recategorisation of information from Table 12.29 to Table 12.30 for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2,3)	2A (4,5,6,21)	2B (7,8)	2C (9,10)	2D (11,12,13,19)	2E (14,19,20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
50101	*	*	*			*		*		Top 10%
50102	-	-	-	-	-	*		*		
50103	*	-	-	-		*				
50104	*	-	-		-	- *				
50105										left class
50106	**	**	*	*	*	*				Top 10%
50107	**	**	**			*				Top 5%
50108	*	-	-		-	-				
50109	*	*	-		-	*				
50110	**	**	**	*	*	**		*	SPAT/ GEOM	Top 5%
50111	**	**	**	*		*		*	HARMONIC	Top 5%
50112	*	**	**	*	*	*		*		Top 10%
50113	*	-	-			*	-	*		
50114	*	-	-			-				
50115	-	-	-			-				
50116	*	*	-		-	-		-		
50117	*	*	-			-	*	*		
50118	*	*	-			-				
50119	-	-	-			*	-	*		
50120	*	**	**		*	*		*		Top 10%
50121	*	**	*	*	*	**		*		Top 5%
50122	-	-	-		-	-				
50123	**	**	**		*	*	*			Top 5%
50124	*	**	*			*	-			Top 10%
50125	*	*	-			*		-	10% spatially	
50126	-	**	*			*	*	**		Top 10%
50127	**	**	**	**		*	*		SPAT/ GEOM	Top 5%
50128	**	**	**		**	*			SPAT/ GEOM	Top 10%
50129	**	**	*	**	**	**	-	**	SPAT/ GEOM	Top 5%
50130	*	*	*			-				
50131	-	-	-	-	-	-	-		SPAT/ GEOM	
50132	-	-	-	-	-	-				
50133	*	*	*		*	*				Top 10%
50134										left class

Class 001

We have seen evidence from previous classes that teachers found difficulty in recording an adequate amount of clues to the pupils' abilities for the identification process for them to depend on their checklist records. We have also seen that adequate clues to ability could be recorded as a result of classroom interaction, but only if steps were taken to spend time with the individual pupils discussing their work and challenging them in appropriate ways.

When the observer was present in the classroom it seemed to be that the teacher would leave him to do most of the work with the pupils during the 'active' phase, so that it was difficult to draw conclusions about how easily the teacher himself could have gathered the clues.

In class 001, however, the teacher worked entirely on her own and the research officer met with her when appropriate to discuss progress, advise on individual checklist items and techniques of clue gathering. We can consider Table 12.32 as the accumulation of half a term of 'semi-active' work and half a term of where the technique was much the same as in the 'active' phase of observation for other classes, but this time carried out by the teacher entirely. At half term there were inadequate clues for proper identification, but we can see from Table 12.32 that this was adequately developed to Tables 12.32 and 12.33 after the 'active' phase. This confirms that a teacher can carry out the activity when the observer is not present. There are no minus signs on Tables 12.32 and 12.33. The interpretation for these tables only, is that asterisks record definite strengths, no asterisks indicate weaknesses.

The teacher in question said that she was 'not 100%' satisfied that further clues could not be found but it was a sound basis for the identification to proceed. It is quite acceptable that a teacher should wish to continue to update the records and this is the strong point of this method of identification, that the teacher constantly reviews the childrens abilities in the light of recent developments of ability and makes provision in the light of the strengths and weaknesses on the profile of abilities.

TABLE 12.10

Checklist levels recorded by teacher
(One Term)

CHECKLIST CODES

Pupil codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
00101	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*		
00102	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*		
00103																											
00104				*		*								*													
00105			*																								
00106	*	*	*	* *	*	*	*	*	* *	*	*	*	* *	*	*	*		*	*	*	*	*	*	*	*	*	*
00107						*	*		*								*			*			*		*		
00108	* *	*	* *	* *	*	*	*	*	*	*	*	*	* *	* *	*	*	*	*	*	*	*	*	*	*	*		*
00109	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		*		*	*	*	*	*	*	*		
00110																											
00111						*			*																		
00112	*		*	*		*			*		*	*						*					*				
00113	*	*	*		*	*	*	*	*	*		*								*			*				
00114				*															*								
00115				*		*			*					*					*								
00116																											
00117	*	*				*			*										*			*	*				
00118	*			*	*	*	*	*	*			*															
00119	*			*	*	*			*		*																
00120	*		*		*	*	*		*	*		*	*				*		* *	*	*	*		*	*	*	
00121																											
00122	*		*	*	*	*	*		*	*	*	*		*	*	*	* *		* *	*	*	*	*	* *	*		
00123				*		*			*										*								
00124				*		*	*		*			*					*		*		*	*	*	*	*		*
00125			*	*		*			*			*					*		*								
00126																											
00127			*			*			*																		
00128		*		*			*	*					*							*							
00129	*	*	*	* *	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
00130	*				*	*			*			*					*		*				*	*			
00131	*	*	*		*	*	*	*	*	*		*	*			*	*		*								
00132					*	*			*										*								
00133					*	*			*										*								
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.33

Recategorisation of information from Table 12.32 to Table 12.33

for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2, 3)	2A (4,5, 6,21)	2B (7,8)	2C (9,10)	2D (11,12, 13,19)	2E (14, 19, 20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
00101	*	* *	*	*	*	*	*	*		TOP 5%
00102	*	* *	*	*	*	*	*	*		TOP 5%
00103										
00104		*			*					
00105	*									
00106	*	* *	* *	* *	*	*	*	*	[illegible]	TOP 5%
00107		*	*		*		*			
00108	* *	* *	*	* *	* *	*	*	*	[illegible]	TOP 5%
00109	*	* *	*	*	*	*	*	*		TOP 5%
00110										
00111		*	*							
00112	*	*	*	*		*		*		TOP 10%
00113	*	* *	*	*			*			TOP 10%
00114		*				*				
00115		*	*		*	*				
00116										
00117	*	*	*			*		*		TOP 10%
00118	*	* *	*	*						TOP 10%
00119	*	*	*							
00120	*	*	*	*	*	* *	*	*	[illegible]	TOP 10%
00121										
00122	*	*	*	*	* *	* *	*	*		TOP 10%
00123		*	*			*				
00124		*	*	*	*	*		*	[illegible]	TOP 10%
00125	*	*	*	*	*	*				TOP 10%
00126										
00127	*	*	*							
00128	*	* *		*			*			TOP 10%
00129	*	* *	*	*	*	*	*	*	[illegible]	TOP 5%
00130	*	*	*	*	*	*				TOP 10%
00131	*	* *	*	*	*	*				TOP 10%
00132		*	*			*				
00133		*	*			*				

Conclusions

This close study of a sample of classrooms has enabled a number of very important points to emerge regarding the use of checklists in mathematics classrooms. We have found particular points from particular kinds of teaching environment, but our most telling conclusions are of a general nature. We recall that the original sample of schools were representative of the County of Oxfordshire and that in general terms the effectiveness of checklists from the first part of our work was very small throughout all of the schools in the sample. The fact that we have isolated some general points from the second part of our work from a sample of these schools leads us to suggest that these discoveries are applicable to many more schools of the kinds from which our sample was drawn.

We have been unable to look in detail at more than one subject but the classroom constraints under which mathematics is taught are similar for other subjects so that many of our conclusions may be generalisable to these subjects.

We will now bring together the main findings of this chapter.

a) The rate at which clues to pupils' ability occur in the classroom depends on the teaching style adopted. One extreme is the style which allows for a large amount of class co-operation and discussion to develop a topic. This tends to absorb a large amount of class time on whole group work. The other extreme is where there is very little whole class activity and this time is spent quite formally introducing a topic with minimal pupil participation. The majority of the time is spent on individual work from text book or worksheet.

The former method encourages a wide variety of clues to ability to emerge, but the teacher finds difficulty in recording these clues, being pre-occupied with the development of the lesson. For certain pupils who fully co-operate the challenges can be at a high level.

The latter method encourages only a narrow range of clues to emerge, and challenges are at a low level. This method, however, allows more time for a teacher to respond to individual pupil's needs during the time available for quiet work. Clues to ability can emerge from this interaction. If, however, a teacher only responds to the pupils in difficulty weaknesses, rather than strengths are observed.

In none of the classes that were observed were the teachers able to record adequate clues to ability for reliable identifications to be based on them, neither was the observer able to gather enough information from passive observations though in all cases he was able to record much more about the pupils' abilities than was the teacher.

b) In every class that was studied it was possible to achieve a full profile judgement on all pupils. In each class that was observed the recording of additional clues became the responsibility of the observer, though the means of increasing pupil response was a co-operative exercise between teacher and observer. Evidence from one school, however, demonstrates that it is possible for a teacher, herself, to keep such records and to increase the availability of the full range of clues to ability.

c) There were two methods of increasing the knowledge of pupils' abilities. First the teacher could scan the records kept and see where knowledge of particular pupils was lacking, then, during an appropriate part of the lesson, engage that pupil in mathematical activity that would reveal knowledge of a particular aptitude. Confirmation of what appeared as strengths or weaknesses could be gradually sought over further lessons until a confident profile emerged. Second, if there were certain checklist items that were not filled in for a large proportion of pupils, the teacher could assess if the pupils were being adequately challenged, and exercised, on this particular aptitude. Lessons could be deliberately modified to bring in such challenges. These two approaches were crucial to both adequate provision and identification. In no school was it possible to properly identify pupils without engaging in the former activity in a most deliberate way. The necessity for the latter activity must be of major concern if able pupils are to be fully nurtured in the subject and in every class that was observed much of the work was predominantly routine and produced a very limited challenge.

d) Workbook scrutiny produced a very limited range of confident clues. This must be, again, compared with the wealth of knowledge that can be gained by discussing the work with the pupil rather than trying to make judgements just from workbooks. The manner of assessment that was used to classify the knowledge of pupils ability from the workbooks, was, of course, different from the normal grading and recording carried out by a teacher. The implications however, is that assessments, of whatever kind, based on workbook knowledge must be based on this narrow range of clues. If we add to these workbook clues the minimal number of clues that teachers were able to record from within the classroom, in the normal run of the lessons, we begin to understand how, when presented with decisions on who ranks as most able, the limited knowledge could be misleading and biases of numerous kinds begin to have affect on the kinds of pupil rated as highly able.

e) The ability of an observer to record clues to ability better than a teacher whose attention was on many other matters is an important point which could be exploited. A strategy for identification could well be to invite an observer into the classroom for a sufficient number of lessons to plot profiles of pupils' abilities for the purposes of identification. A small number of lessons may suffice to show the teacher where further pupil/teacher interaction is needed in order that he might complete the profiles alone. This teamwork could lead not only to better identification, but also to a review of the adequacy of provision.

f) Certain checklist items were underchallenged, and hence hardly observed, even when positive steps were taken to improve the variety of clues to ability. Here, again, we have conclusions which emerge from the relationship between identification and provision.

Pupils were rarely given problems of the 6c variety - problems that were of a nature requiring exploitation of a wide range of knowledge or mathematical insight. More than just a challenge from a set of miscellaneous exercises is required, though very little time was generally given for even this. We are thinking of problems for which there is no obvious prescription. Too often the work, even for the most able, was repetition of routine questions (6a) or slight deviations (6b). Out of these '6c' problems we would expect deeper challenge of many of the checklist items: handling of large amounts of information, from which the relevant information has to be isolated and often where there is inadequate information (items 2 and 3); the production of logical proofs, (items 4,5,6); the exercising of the aptitude for flexibility in devising and understanding proofs, (items 12,13); the seeking of a variety of solutions from which an elegant one emerges (items 14,15).

One particular item that was tentatively recorded, at best, was 'mathematical cast of mind'. It may be that at 3rd year level clues could not be strong enough to detect particular bents (spatial or analytic). Inevitably, therefore, clues that were picked up eminated from a wide range of mathematical topics. It was possible to judge pupils in relative terms as a consequence of the level of work presented, mainly by extending the challenge within these '6a and 6b' type of problems with individual pupils.

Perhaps the most appropriate suggestion is that when pupils have displayed their profile of strengths in the identification process, in relation to the normal classroom challenges, the sorts of challenge that these pupils are given should be reassessed so that they might exercise these strengths and be allowed to escape from constant and limiting repetition of routine work. This may not be seen as the most appropriate course for others in the same class who have weaknesses on the profile, though a gradual strengthening of these weaknesses may be considered necessary.

g) We have shown that it is possible to make valid identifications of pupils in the classroom which are entirely a result of checklist usage. Appendix 7 shows how the checklist was validated and also includes the follow-up work that demonstrates the reliability of the conclusions reached. Most of our previous work was concerned with an identification to be measured against test prediction for O-level performance. In this checklist work, and because of the validity of the checklist, we have escaped from the O-level criterion. We have made valid judgements of present mathematical ability, rather than potential for the future.

The consequence of this is that we can be confident in our ability to monitor the development of pupils, based on what we are at present observing in them and modify our judgements of needs and abilities from week to week. Our provision can be appropriate to the broad spectrum and we are freed from the arguments associated with any narrowness of identification from tests or for an examination such as O-level. Indeed we can have the broad, rather than the narrow challenges of examinations such as O-level fully in mind for the appropriate pupils because of a full knowledge of them and the spectrum of abilities.

h) The effort required for proper identification must not be underestimated. We have seen how positive steps must be taken to work with individual pupils and this must be done over a considerable number of lessons; we cannot expect to set ourselves a few lessons to complete the task once and for all. The most appropriate interpretation of this is that it should affect the teaching style once and for all. We could even be as bold as to suggest that the effort put into this classroom activity - where the pupils workbook will also be under some scrutiny as a basis of the interaction - could be offset by a lesser need for routine marking of the traditional kind. With limited time available for all of the activities that a teacher considers appropriate, minimised out of class workbook scrutiny would enable lessons to be prepared in an appropriate fashion to meet the variety of individual needs of pupils across the mathematical spectrum.

i) Identification was concerned with the top 10% and 5%, relative to the class norms, but it was clear, from the knowledge accumulated, when a pupil of high calibre relative to national norms had been discovered.

Some conclusions based on the prediction equation

We discussed, above, how we have escaped from the O-level criterion in forming our conclusions concerning checklist identification in mathematics. We will, however, give some results similar to previous ones related to earlier work.

Tables 12.34 and 12.35 can be compared with Tables 6.8 and 6.9. We recall that because of test error we would expect a certain number of mismatches between test prediction and final membership of the 10% group on actual O-level score. The 'checklist' row contains the total number of pupils in each score band for the school identifications reported in this section. These are to be compared with the 'test forecast' from the larger pooled sample from the first years work and the 'Round 1' identification of the first year which was completely without checklist assistance.

The purpose is to compare the distribution of numbers identified and not by checklist within each band, first with the expected and 'allowable' numbers, and second with the distribution that came from non checklist usage. The comparisons are 1 row Chi-squared tests of goodness of fit similar to those in chapter six.

TABLE 12.34

NOS. OF PUPILS NOMINATED BY TEACHERS AND NOS. FORCAST BECAUSE OF TEST ERROR WHOSE PREDICTED SCORE LAY IN EACH BAND OF WIDTH ½ A STANDARD ERROR BASED ON A DATUM OF THE 50th PERCENTILE

		BELOW 90th PERCENTILE				ABOVE 90th PERCENTILE		
		1½-2	1-1½	½-1	0-½	0-½	½-1	1-2½
	CHECKLIST	1	4	5	10	22	12	19
MATHS	TEST FORCAST	6	15	30	40	51	40	28
	ROUND 1 (YR1)	3	8	27	45	46	45	37

TABLE 12.35

NOS. OF PUPILS NOT NOMINATED BY TEACHERS AND NOS. FORCAST BECAUSE OF TEST ERROR WHOSE PREDICTED SCORE LAY IN EACH BAND OF WIDTH ½ A STANDARD ERROR BASED ON A DATUM OF THE 50th PERCENTILE

		BELOW 90th PERCENTILE				ABOVE 90th PERCENTILE		
		1½-2	1-1½	½-1	0-½	0-½	½-1	1-2½
	CHECKLIST	113	100	75	38	17	4	0
MATHS	TEST FORCAST	349	275	212	119	62	22	8
	ROUND 1 (YR1)	324	310	229	134	56	19	8

Table 12.36 lists the chi-squared values and significances that emerged.

TABLE 12.36

CHI-SQUARED VALUES AND SIGNIFICANCES FOR GOODNESS OF FIT TEST APPLIED TO EACH BAND IN TABLES 12.34 and 12.35

		BELOW 90th PERCENTILE				ABOVE 90th PERCENTILE		
		1½-2	1-1½	½-1	0-½	0-½	½-1	1-2½[1]
CHECKLIST V TEST FORECAST	χ^2		1.93	1.57	0.57	0.97	0.17	3.16
	SIG		N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
CHECKLIST V ROUND 1	χ^2		0.37	2.78	0.48	2.00	0.57	5.07
	SIG		N.S.	N.S.	N.S.	N.S.	N.S.	5%

---

[1] Small numbers forced us to combine cells ½-1 with 1-2½ for the final comparison.

The single significance in this table indicates a slight improvement of identification when checklists usage is compared with non checklist usage. This occurred in the highest percentile bands. Otherwise conclusions are much the same from this quantitative work as we found in Chapter Six. However, we know that teachers would certainly have had little help from their checklists in the earlier work reported in Chapter Six, whereas this time we are confident in the particular choices they made, so both qualitatively and quantitavely checklists have found pupils with whom test prediction for O-level potential would be in agreement.

The higher numbers of pupils who were nominated in the top percentile score band compared with identification without checklist is confirmed if we count the numbers of A Group pupils who were nominated. Indeed every one of these pupils high scoring pupils were identified by checklist.

In view of the fact that we are confident that the checklist identified pupils have been nominated because of mathematical aptitudes we need not look at test based characteristics as we did before. One useful comparison, however, will be included. This concerns the numbers of girls and boys who were selected. Table 12.37 gives these numbers.

TABLE 12.37

NOS. OF GIRLS AND BOYS NOMINATED FROM CHECKLIST
BAND IDENTIFICATION

	1	TOP 10%	TOP 5%
BOYS	1	45	16
GIRLS	1	29	12

In order to see if these differences are meaningful we carried out goodness of fit chi-squared tests to compare with an equal distribution of boys and girls. Neither the top 10% numbers nor the top 5% numbers showed significant difference between numbers of boys and girls who were nominated.

Finally, we can refer to the list of constructs elicited from the teacher in School OO that followed the checklist work. These are given in Appendix 8. There we have studied the ability of teachers to describe pupils and have found that they only refer to a limited range of abilities and attitudes, mostly of a general nature rather than specific aptitudes such as appear on the checklists. The teacher in Class OO1 had recorded valid checklist levels on the broad spectrum for the class she was teaching, yet we were still only able to elicit the limited range of constructs which are associated with insuffient knowledge for proper identification.

It was not possible to pursue this investigation in a large number of schools so we can only say that this limited evidence from one school suggests that it was not possible for a teacher to remember the details concerning the pupils when that teacher was interviewed, and that identification by checklist was only possible because of the adequate records that were kept. This important point re-emphasises the need for a teacher to take deliberate and methodical steps if the checklist based identification is to be effective.

Note for Chapter Twelve

1) In this class as in all of the others the lessons ranged considerably over different sorts of mathematical topics. The clues are an accumulation of all kinds, algebraic, geometric and logical. Insufficient clues would have been gathered from any one branch of the subject.

CHAPTER THIRTEEN

Conclusions and Recomendations

13.1 Introduction

The reader who has worked through the preceding chapters will have had his own particular environment and requirements in mind and will, we hope, have been able to draw many conclusions which are relevant to his own situation. It would be impossible for us to draw together all of these specific conclusions here, and we will make no attempt to do so. Nevertheless, we feel that there are some more general conclusions which can be drawn from our study. As the work has taken place in a sample of 'normal' schools in Oxfordshire, we feel that these general conclusions can be expected to have relevance for a large number of comprehensive schools which operate under similar constraints and with similar expectations.

We will first list these general conclusions and then go on to discuss some of their implications.

13.2 Summary of general conclusions

1. Test/teacher agreement (Ch.3): In each of the four subjects that we have studied, we have discovered a greater level of agreement between test-based and teacher-based nomination of pupils who were in the top 10% and top 5% of ability, than previous research might have led us to expect. This previous research, however, was concerned with identification for general ability, whereas our study was concerned with identification within a given subject. In each subject the teachers and the tests were asked to identify the pupils with the greatest potential for O-level performance.

2. English and Mathematics (Ch.3): The level of mismatch between test-based and teacher-based nomination in English and mathematics was at a sufficiently low level that there was no evidence to suggest that teachers were any less effective than tests in their nomination of able pupils. Nevertheless there were a number of high scoring pupils omitted from the teachers lists and a number of low scoring pupils included on the lists.

3. French and Physics (Ch.3): In French and physics there was sufficient mismatch between test-based and teacher-based nomination to indicate actual errors of judgement on the point of the teachers. Too many high scoring pupils were missed from the teachers' lists and too many low scoring ones were included by teachers, for the mismatch to be accounted for entirely by the error that was present in the test-based assessment of O-level potential.

4. School differences (Chs.3 & 4): There was no detectable difference between schools of different types (large/small, single sex/mixed, 11+/13+) in relation to the levels of teacher/test agreement.

5. Bias in teachers judgement (Ch.4): There was strong evidence associated with test scores that bias was present in teachers' judgements so that high scores on inappropriate criteria often characterised a teacher's list of pupils.

6. Teacher differences (Ch.4): The sorts of biases that existed were associated more with different teachers than with different school types.

7. Social class and sex bias (Ch.4): Social class and sex differences were not the dominant biases though biases of this kind could be detected in some of the teachers' lists.

8. Physics general ability bias (Ch.4): There was a strong indication that physics teachers identified pupils with high general ability rather than high ability on other criteria which were more relevant to physics ability. This effect was most noticeable when identifications were made early in the year. It was not so evident for identifications made later in the year than when the teachers had taught the pupils for longer.

9. French favourable attitudes (Ch.4): There was strong evidence that children who had favourable attitudes to the subject were selected for the French lists.

10. Changes on the teachers lists (Ch.6): The number of children who were replaced by others on the teachers' lists, on the second round of identification, was of a sufficient level in all subjects to indicate

that there had been uncertainty in one or other of the identifications. This was particularly so in the case of physics.

11. *Checklists* (Ch.8): There were strong indications that checklists were not effective instruments in guiding the identification process when teachers were asked to fill them in for each pupil, drawing on recent observations and memories of the pupils' abilities. There was strong evidence that the teachers did not have the detailed knowledge of the pupils on the broad spectrum of subject aptitudes that would be required for this exercise.

12. *Identification for provision* (Chs.7 & 11, App.7 & 8): One implication was that, though teacher identification of those pupils who could obtain a good O-level pass in the subject (particularly in mathematics, and English) may have been reasonably successful, the mapping, by teachers, of pupils' individual strengths and weaknesses *within the subject* was not very detailed (except in English). It is clear that without this detailed knowledge the structure of the teaching and of the challenges that are set to the pupils in the subject cannot take full account of these strengths and weaknesses.

13. *Validity of checklist* (Ch.12, App.7): Agreement as to what might constitute a valid checklist was not always easy to achieve, particularly in English and French. This indicated some uncertainty between teachers as a whole as to the specific aptitudes which constitute and should be developed through the teaching of their subject. In the case of mathematics it was possible to compile a valid checklist because previous research in the subject had clarified the nature of these specific aptitudes. Even though we have discovered that there is considerable difficulty associated with the use of checklists, it must remain a priority for teachers to agree on the profile of aptitudes that make up their subject whether it be for purposes of guiding identification or structuring provision. A clear need for the sort of research in other subjects, that provided the background to the mathematics checklist, has been demonstrated in our study.

14. *Checklists used in another way* (Ch.11): Similar conclusions regarding the ineffectiveness of checklists was discovered in English, French and physics when teachers adopted the seemingly more appropriate strategy

of assembling the clues to ability over several months.

15. Mathematics classroom (Ch.12): The lack of sufficient clues to ability, that would enable a subject specific profile to be built up for each child, was confirmed, by observation of the mathematics classroom. It was discovered that sufficient clues were not observable in the children, nor, of those clues available, were adequate records kept by the teacher, the latter being due to the preoccupation with other classroom activities.

16. An effective method of checklist use (Ch.12): It was discovered to be possible, however, to gather sufficient clues to ability if the teacher took deliberate steps to do so. This was accomplished by the teacher spending more time with individuals than would occur naturally. Appropriate challenges could be presented to an individual from which a range of clues could emerge. It was also possible to observe more clues to ability if the work that was given to the class was varied, in order to challenge the pupils more broadly.

17. Workbooks (Ch.12 App.7): A very limited range of clues to ability could be extracted from the workbooks of children in mathematics, so that appropriate classroom activity was the main vehicle to valid identification.

18. Checklist as an effective instrument (Ch.12): The conclusion for mathematics was that the checklist was a reliable and useable instrument, but only if a teacher made deliberate attempts to gather clues. Even when this was done, however, memory of the profile of abilities for each child was not possible when the records were not available, which emphasised the need to keep such records. The 'off the cuff' memories that teachers had of their pupils were mainly connected with broad generalisations about subject strengths and opinions about their aptitude and motivation.

19. Characteristics of pupils selected as more able (Chs.10 & 11): Insights were gained in to those characteristics of pupils to which teachers were sensitive and into the role that assessments of these characteristics played as teachers made their overall judgement of pupils' potential in each subject. In English, teachers were often able to rate their pupils on a wide range of sometimes quite

sophisticated English specific abilities. In physics, teachers displayed a tendency to make assessments of pupils' general ability and of their motivation and personality. They did not seem to rate pupils on a wide range of physics specific aptitudes. Some evidence was assembled to suggest that physics teachers' ratings of pupils' on various motivational and personality dimensions were important in contributing to the inaccuracy in teacher judgement of overall potential in this subject.

20. Pupils of high ability in more than one subject (Ch.1): We have brought to light strong evidence that the number of pupils who have high ability in a large number of subjects is quite small. There is agreement between test-based and teacher-based assessment on this issue. This strongly supports a subject-specific focus for identification of, and provision for, the more able and for any system of grouping applied to pupils as a whole.

13.3 Identification by test or by teacher?

The main sort of tests that we are considering here are the standardised tests of general or specific abilities that are widely available. We have shown (in Appendix 8) that appropriate combinations of such tests can give useful evidence which relates to pupils' abilities: evidence that can be helpful in connection with the identification process. However, only extreme pessimism about the ability of teachers to know the subject specific strengths and weaknesses of their pupils, would lead us to suggest that the use of such tests in isolation could provide the complete answer to identification. We rehearse several reasons for this view below.

First, in selecting which tests should be used, there is a risk that teachers will place too much weight on familiar IQ tests, and will undervalue wider batteries of tests which, though less familiar, are more time consuming in administration and marking, and more complex in interpretation, can be more effective as indicators of subject specific ability.

Secondly, there is a tendency for a 'once-and-for-all' judgement to be made about pupils when tests are used. Good test programmes will always include repeated measures so that changes in ability can be mapped. However the commitment of time and considerable expense that would be implied by repeated testing (especially repeated testing with a large test battery) may lead teachers seriously to reduce the frequency of testing.

Thirdly, it is understandably tempting, when decisions *have* to be made about the selection of individual pupils for some activity, to overlook the problem of the error associated with the test, and to make what are essentially false distinctions between pupils who score 120 and those who score 119. Indeed if one were to base identification *entirely* on tests, one would have no alternative but to do so.

Fourthly, however wide the test battery, there is much that it will not measure about pupils that would be relevant to accurate prediction of their future performance.

In this context it is fortunate that we have been able to show, in all four of our subjects, that the accuracy of teachers' judgements was sufficiently high for the teachers' own contributions to the assessment of pupils to be regarded as of considerable relevance. In Chapter One, we drew attention to the work of Pegnato and Birch, who had found a 45% level of agreement between teacher and test - in their case an I.Q. - test which we can take as representative of the findings of other researchers too. We have taken the analysis on much further and considered the purpose for which the test might be given. In considering that purpose - in our case prediction for O-level score test error becomes an important part of the analysis and the tests scores themselves cease to be thought of in the same way. Our work, from this point differed from that of Pegnato and Birch and of others, who took no account whatsoever of test error. It is as a consequence of our interpretation of test scores that we have obtained the more telling conclusions about teacher accuracy. It also remains true, however, that levels of teacher/test agreement are, in general, much better than the 45% figure even when no account is taken of test error. This, we would suggest, is because of the subject specific rather than I.Q. focus of our work.

However, we also have enough evidence to suggest that teachers' estimates of ability can be affected by seemingly inappropriate pupil characteristics. This is especially so in some subjects. Also there is a tendency for teachers to refer to general abilities rather than to a detailed profile of subject specific abilities. As a consequence, some inaccuracies of judgement are likely to be present. This implies that we cannot blithely discard the sort of evidence that appropriate psychological tests provide and rely solely on 'off the cuff' teacher assessment. Even with some test evidence, however, there cannot be any doubt that we must be concerned about the gap in the detailed knowledge that teachers have of their pupils when a conscious and deliberate effort is not made to know them better. Though there is no particular factor that consistently subject teachers individual biases do exist for each teacher, sometimes connected with personality, behaviour, sex or social class but usually connected with inappropriate academic abilities. Whether or not tests are used as an aid to identification this gap of knowledge should be filled. The most important thing in assessing children in a specific subject is that the teacher must know the conceptual structure that subject in detail, with all its branches, networks and hierarchies. Sections of this knowledge are of special importance to children at different stages so that detailed knowledge of the children's strengths and weaknesses within a subject will allow the provision to be designed to fill the needs.

A possible strategy for identification is that tests could provide some evidence, and that the accumulation of teacher observations could add to this evidence day by day, neither of the sources being the sole means of identification. This is a plausable solution, but further research is called for to bring to light the best ways of using test information and teacher information together to form some overall judgement. We need to establish what it is that one approach adds to the other, and how the most appropriate parts from each can be combined. In the absence of such detailed research, perhaps a sensible starting point for this approach would be to obtain test results and independant teacher nominations, and to look critically at pupils who appeared on one list but not the other. The nature of the 'critical look' needs to be explored but may perhaps involve a diagnostic interview with the pupil. Ideally, this 'plausible starting point' would, itself, be the subject of evaluative study. This is an interesting field for exploration, but in the light of other insights we have gained it is just one of a number of options and possibly not the favourite.

13.4 Use of checklists

We have demonstrated that checklists have not been the effective instruments that many have supposed. One problem is associated with the accumulation of items for the list which are valid as indicators of high ability in the subject, but which are also acceptable to the teacher and observable in the classroom. The other problem is the difficulty of maintaining adequate records of a pupil's abilities as they are displayed. The consequence is that the job is just too difficult to do unless it becomes one of the main priorities.

To suggest that it should be a priority may provoke a sharp initial reaction, but if we consider that a teacher can only meet the needs of any individual, very able or not, by having a clear knowledge of that individual's full profile of aptitudes it becomes apparent that full profile knowledge should be a priority for all pupils. The link between identification and provision is quite apparent, for clearly such a full profile can only be assembled if the aptitude is challenged by the work that is set for pupils.

We have seen, from the mathematics study, that if a teacher makes a deliberate effort to work closely with each individual pupil during class time a full checklist profile can be built up. As the mathematics checklist items were carefully validated, the resulting profiles could be used as the sole means of identification. The fact that the teacher must use a checklist quite deliberately should not be underestimated, but the resulting accumulation of very detailed knowledge about individual pupils is ample reward, especially as this knowledge can be used for more than the original purpose of identification. For example one further use of the checklist approach arises from the fact that the completed instrument not only provides detailed information about individual pupils, but also a record of the range of challenges that normal lessons present to the class as a whole. If it is interpreted in this way the checklist becomes a general curriculum development device as well as a means of identifying the more able pupil. Another further use is that it enables the teacher to design appropriate provision for individual able pupils and to monitor their performance on the additional work that is presented.

If this approach is to be generalised to subjects other than mathematics where our own checklist has been carefully validated, validation of the corresponding checklist items would be needed. Perhaps the mathematics study could be taken as an indication of how such a validation might be carried out.

13.5 A Second Observer

We discovered from the work in mathematics how much more a second teacher could record concerning the individual abilities of a class of the pupils in a class, than was possible for the teacher who was at the front of the class engaged in all of the activities involved in delivering a lesson. This was mainly because of the single minded approach that the second observer could adopt.

We would suggest that the boost to checklist profile construction that the use of a second observer could give to a class teacher, is considerable. The effort that is required for the provision of an extra teacher for a sample of lessons could be compared with the effort that could be saved if it were no longer necessary to administer and mark psychological tests. We would argue that the use of a second observer is potentially a very interesting basis for the identification of the more able and that it should be given careful consideration.

13.8 Identification through provision for a larger group

Often our identification of able pupils is for the purpose of supplying the pupils with extra provision. Often too, we are constrained by numbers of pupils that we can handle in such a provision programme, especially if it takes the form of tutorial teaching or summer school activity or the like. If this is so we must seek a method that gives an accurate ident ification of the appropriate number of pupils such as those that are the subject of this chapter. However, when there is more flexibility over numbers, we could adopt the strategy of asking teachers to nominate the pupils without any test or checklist support, and without the in-service training that we will discuss later, and we could agree to allow in more than the number for whom the provision was designed. This allows in any pupils for whom the teacher feels some uncertainty. The level of provision would still be for the same high level of ability so that when the pupils were partaking of the provision, the teacher could begin to pick out those who could not reach the levels required to continue with the provision programme. If the way in or out of the enrichment group was easy, there there would be little repercussion for the child who needed to drop out.

It may be that, if teachers are asked to nominate a small number of pupils for a summer school, as an alternative to one of the procedures discussed above, they might begin by choosing a larger group for some extra provision in school and then make their final selection from this group in the way suggested above.

We would suggest that this method is in line with our findings and applies to whatever level the provision is aimed. For example when we asked teachers to nominate their top 10%, most of the top 5% on test mark were included, so that if our provision were for the top 5%, allowing to the top 10% in the initial stages of the provision could enable us to find the right pupils.

The selection could be entirely dependant on teacher nomination, and might rely on just the teacher's day to day knowledge of the pupils.

11.7 In-service training

In our work we have discovered that teachers were often uncertain about what can be expected of, and what should be provided for an able child in their subject. More specifically we have found that their ability to recognise able children was influenced by pupil characteristics that often seemed to be rather inappropriate. Since these influences were, no doubt predominantly subconscious, it is very difficult for a teacher to gain the feedback which is necessary to enable him to make adjustments to the things which he allows to influence his judgement. We would suggest that our work provides a basis on which teachers could be helped to refine their understanding in this area. It is interesting to speculate that the characteristics of teacher selected children, which were revealed in Chapters Four and Eleven, could be used as a starting point for a series of discussions with teachers about identification.

Whilst strongly advocating this in-service training approach, we would also wish to ensure that in-service initiatives were evaluated. The risk, if this is not done, is that one apparently sensible, but previously untested answer to the question of how teachers' accuracy might be improved, (namely checklists) may be replaced by another (namely in-service training). Nevertheless, without time set aside for proper training it is unlikely that the teacher would have time to reflect on the major issues that are associated with both identification of and provision for able children.

In a similar way, it is possible that the profiles of lessons provided in Chapter Twelve could be used as the basis for discussions about classroom provision for the more able.

13.8 Pupil grouping

We can only reflect briefly on this issue, but would be wrong to neglect it completely.

There are a number of ways in which pupils can be grouped for teaching purposes. Whenever decisions are made connected with ability grouping, an identification strategy has to be employed.

The effects of mistaken placement of pupils of high ability in low sets is well known. We have studied the ability of teachers to nominate children of high ability, and found that in all subjects, but much more strongly in some than in others, errors are made.

Some of the errors were made before the third form teacher drew up his lists so that some pupils who were assessed as having high ability by tests were not even considered by him. The system had already condemned these pupils by placing them in low sets or streams. It would seem that if the needs of all of the more able are to be best served, children should at least begin their early secondary school career in mixed ability groups and that the transition from this type of grouping, to another based upon ability, should occur only when sufficient information has been assembled to enable this to be done in a subject specific way. A wide range of (not just I.Q.), tests, teachers opinions guided by checklists, and teachers opinions honed by the inservice training that we have suggested, may all have a part to play in building up the package of information upon which this subject specific setting might be based.

In putting forward this view that a mixed ability start would be advantageous to able pupils it is interesting to note that the Banbury Grouping Enquiry (Postlethwaite and Denton, 1978) found that able pupils who spent their first year in mixed ability groups gained slightly better examination results than those who had spent their first year in streamed groups.

Furthermore, the evidence that we have provided regarding the relatively small number of pupils who were of high ability in a large number of subjects provides a strong warning to those who favour a rigid policy of streaming across several subjects on the basis of general ability alone.

There are, therefore, organisational implications in some of our findings, as well as implications for individual teachers who wish to improve their own ability to select able children from the classes that they teach. If we do not take up both of the issues in a positive way we may fall well short of our duty to the able children in our care.

APPENDIX 1

Test battery and prediction equations

For the convenience of the reader we will list the test scales that were administered to our samples of pupils and the prediction equations that were derived from them, on one fold-out page that can be used alongside the relevant section of the report.

The same 3rd year test data was available on two groups of pupils who took O-level examinations in their 5th year in 1977 and 1978. It was from the test scores, and the subsequent O-level results of these pupils, that the prediction equations were established.

For each subject, the prediction equation was derived by multiple regression analysis. Very briefly the procedure was that a matrix was established showing the correlations of each test score with every other test score, and the correlation of each test score with O-level result. The matrix was inspected to discover which test scale correlated most highly with O-level performance - and, therefore, which test scale was the most effective predictor. This scale became the first scale in the prediction equation. Inspection of the matrix continued to identify other scales which could be combined with the first to improve the prediction ability of the equation, and these scales, if found, were then included with appropriate weights. The procedure, which was automatically carried out by computer, did not rely on any subjective judgement of which test scale *should* be a good predictor for O-level, but simply indicated which combination of scales was *in effect* the best predictor.

TABLE 1. THE TEST BATTERY

Differential Aptitudes[1]	VR	Verbal Reasoning
	NA	Numerical Ability
	VN	VR+NA (General Scholastic Ability Scale)
	AR	Abstract Reasoning
	CSA	Clerical Speed and Accuracy
	MR	Mechanical Reasoning
	SR	Space Relations
	SP	Spelling
	LU	Language Usage
Attitudes[2]	A	Attitude to School
	B	Relationship with Teacher
	C	Academic Self Image
	D	Attitude to Class
	E	Social Adjustment
	F	Anxiety in Class
	G	Importance of Doing Well
	I	'Other' Image of Class
Creativity[3]	Flu	Fluency
	FLEX	Flexibility
	ORIG	Originality
Other Variables	SEX	Boy or Girl
	NT	Neatness[4]
	SC	Social Class

TABLE 2
THE PREDICTION EQUATIONS

SUBJECT	EQUATION FOR CALCULATING PREDICTED SCORE	% VARIANCE	STANDARD ERROR	BEST SINGLE PREDICTOR	% VARIANCE
ENG	0.65 LU + 0.35 SP + 0.66 VR + 2.64 B + 31.26	36	15.18	LU	22
FR (0120)	2.17 LU + 9.44 G - 7.18 A + 1.68 AR - 0.86 SR + 1.04 NA - 82.52	49	19.00	LU	28
PHYS	1.45 MR + 0.54 NV[5] - 0.36 SP - 0.72 CSA 59.05	44	22.61	MR	25
MATHS	2.66 NA + 0.72 MR + 1.53 AR - 6.6 E + 0.79 VR - 21.77	33	30.32	NA	21

NOTES

1. 5th Manual for the Differential Aptitude Tests (Bennett, G.R., Seashore, H.G., Wesman, A.G.) Psychological Corporation 1974.
2. NFER SF7 scales. Attitudes to School.
3. NFER SF4 Free Writing Test.
4. The neatness mark was on a 1 to 5 scale and was assessed by the research officers who were able to standardise the scoring over the whole sample of pupils by giving a neatness score to their test papers.
5. NV = NA x VR.

Derivation of the equations

One problem in our research design was that of deciding upon measures which can be applied to children at age 13, and which can be used to predict for future O-level performance in each of the four chosen subjects. Both from the theoretical and the empirical standpoint, there is evidence to suggest that IQ is not the most appropriate measurement for the assessment of subject specific potential. As discussed in Chapter One Guilford's model of the Intellect and similar multidimensional models of intelligence, provided some theoretical basis for such a view. Empirical support for the view that general ability, as measured by IQ, is not a particularly good predictor of subject specific ability, can also be found in the literature. For example, Vernon[6] has quoted correlations between IQ and School Certificate subject results which range from .51 (Mathematics) and .50 (English), through .38 (History) and .39 (Chemistry) to .12 (Biology) and -.02 (Art). Thus, even in the cases of mathematics and English, only some 26% of the variance in School Certificate results was explained by IQ. A similar picture has also been obtained by the present Project. We have shown that the percentage variance in O-level scores explained by IQ was 15% (English), 14% (French), 22% (Physics) and 35% (Maths). Fuller details are given in Table 3 below.

TABLE 3

RELATIONSHIP BETWEEN O-LEVEL SCORE AND THE GENERAL ABILITY SCALE DH

SUBJECT	CORRELATION	NO. OF PUPILS	% VARIANCE EXPLAINED
ENGLISH	.39	121	15
FRENCH	.37	59	14
PHYSICS	.47	109	22
MATHS	.59	151	35

Previous work[8], and discussion with psychologists at NFER suggested to us that these levels of prediction could be improved upon by using, instead of IQ, a group of tests made up from the Differential Aptitude Test Battery, together with measures of attitude and creativity. We therefore decided to explore the use of this group of tests in detail and to establish appropriate combinations of test scores for each subject which could be used as predictors of O-level performance in those subjects. We were able to achieve this because data from the same group of tests was available on two groups of 13 year olds in two schools. Roughly a hundred of these pupils who were in their 3rd year in 1974 and 1975 had gone on to take O-level examinations in each of the subjects, but no attempt had been made to identify individuals as more able. It was therefore possible, with minimum distortion arising from the well documented relationship between expectation and performance[9], to establish by means of multiple regression analysis, relationships between pupils' test scores at 13 and their subsequent performance in each of the chosen subjects. The resulting prediction equations, when applied to the data from subsequent groups of 13 year olds in this present study, were used to generate pupils' predicted O-level scores. It is these predicted O-level scores which we have interpreted as objective measures of potential. For immediate comparison with Table 3 we summarise in Table 4 the correlations and percentage of explained variance that resulted from the prediction equations.

TABLE 4

RELATIONSHIP BETWEEN O-LEVEL SCORE AND PREDICTED O-LEVEL SCORE CALCULATED FROM THE PREDICTION EQUATION

SUBJECT	CORRELATION	NO. OF PUPILS	% VARIANCE EXPLAINED
ENGLISH	.60	156	36
FRENCH	.70	68	49
PHYSICS	.66	92	44
MATHS	.57	130	33

Extensive use was made of the SPSS (Nie, Hull, Jenkins and Steinbrenner, 1975) programme for Multiple Regression Analysis to seek the best predictors for the four subjects. Early attempts were concerned with both O-level and CSE prediction on a combined scale, but the most confident approach was concerned with prediction for O-level alone because of the problem of combining two sets of examination results into a single scale. Prediction for an Examination Grade was seen as rather coarse so we investigated the possibility of deriving a predictor for actual scores at O-level.

We were fortunate to gain the cooperation of the Oxford Delegacy of Local Examinations who released O-level scores for the two year groups of the two schools for Mathematics, French and English. The Oxford and Cambridge Board also kindly cooperated to release the Nuffield Physics results. All of this was approved by the schools themselves. This enabled us to satisfy more closely the requirement of regression analysis that the scales should be interval, and we now had much more information than could be contained in the pupils'grade alone.

The numbers of pupils involved in the analysis were as follows:

TABLE 5

1977 EXAM

	: TOTAL NO. IN : YEAR GROUP	NOS. ENTERED FOR O-LEVEL ENG	 FR	 PHYS	 MATHS
SCHOOL 11	: 429	82	79	51	70
SCHOOL 07	: 160	43	24	2	40

1978 EXAM

	: TOTAL NO. IN : YEAR GROUP	NOS. ENTERED FOR O-LEVEL ENG	 FR	 PHYS	 MATHS
SCHOOL 11	: 365	61	40	63	62
SCHOOL 07	: 190	55	12	16	25

We were thus concerned with a predictor for examination success within the group of candidates who were entered for O-level. Since the whole spectrum of results were included we were confident that this predictor could be applied to the whole group in our study of the more able. A second caution was concerned with the particular examination for which our equations predict. The two schools, followed the same examination syllabus in all subjects except French, so that we were able to pool results on all but French. Because of the small numbers taking French in one school, the aim of our prediction technique to produce test based top 10% groups in the schools, we have to mind that the equations predict for a particular syllabus, but we are confident in our approach because many schools used these syllabuses and also because the particular syllabuses which we used have much in common with other popular syllabuses, especially as they affect work at third year level.

TABLE 6

EXAMINATION COURSES ON WHICH THE PREDICTIVE EQUATIONS ARE BASED

	I	ENGLISH	FRENCH	MATHS	PHYSICS
SCHOOL 13	I	01	0120	052	Nuffield
SCHOOL 07	I	01	020	052	Nuffield

Separate predictors were established for the 1977 and 1978 entries. The 1978 cohort at the two schools were tested for Differential Aptitude (DAT) in their third year, while the 1977 cohort was tested in their fourth year. In order to have a common level on which to measure, the DAT scores for the 1978 entry were adjusted to take account of this. This was accomplished most easily by using percentile markings. A pupil's score on a particular scale was checked for its percentile rank based on the NFER national norms for fourth year pupils and he was then given the corresponding score for that percentile from the third year national norm table. No adjustment was necessary for attitudes or creativity because pupils in both years did the tests at the same time in their school career, but a small adjustment was necessary to the examination scores obtained in the two year groups. This adjustment was made within each grade in each subject. The range of scores for a particular grade in 1978 was compared with the range for a particular grade in 1977 and the 1977 mark was adjusted by the appropriate amount. Since this adjustment was of the order of 3 or 4 marks in 150, for the majority of cases a simple ratio adjustment within each grade was thought to be adequate. The results for the regression runs on the four subjects on the two year groups separately is shown below.

TABLE 7

PREDICTION EQUATIONS

1978 Entry

SUBJECT	EQUATION FOR CALCULATING PREDICTED	NO. OF CASES	% VARIANCE	STANDARD ERROR
ENG	0.99R + 4.75E + 0.33SF + 3.35L + 32.91	55	19	12.6
FR (0120)	5.15 E + 66.76	22	15	23.52
FR (020)	8.29LP - 172.15	6	36	36.47
MATHS	3.65NA + 1.46MR - 1.30CSA - 9.08B + 17.49	80	41	31.43
PHYS	1.4MR + 1.85 NA - 1.39CSA 466.5	42	37	32.51

1977 Entry

SUBJECT	EQUATION FOR CALCULATING PREDICTED	NO. OF CASES	% VARIANCE	STANDARD ERROR
ENG	1.16LU + 0.64VR + 0.26SF + 2.58R + 17.58	89	53	14.91
FR (0130)	2.44LP - 5.936	45	32	21.36
FR (020)	3VR + 1.33SF - 2.85NA- 67.07	18	77	18.44
MATHS	3.16NA + 0.85NR + 26.33	79	31	29.45
PHYS	1.83MR + 1.91VR - 2.18E -23.15	49	58	23.06

The prediction equations were applied to their own groups and a Spearman Rank Order Correlation was calculated. The equations produced for the 1978 cohort were used to predict scores on the 1977 cohort (using the adjusted data) and correlations produced. The results were encouraging and led us to the final step of combining (adjusted) results for the two groups and finding the overall prediction equations reported in Table 7. The numbers of pupils in the analysis were English 140, French 68, Physics 92 and Mathematics 130. The best single predictor is also shown with the percentage variance that it accounts for in the prediction equation. Table 8 shows the results of the correlation analysis.

TABLE 8

CORRELATIONS OF PREDICTED RESULTS WITH ACTUAL RESULTS

SOURCE OF EQUATION	CORRELATION ON YEAR GROUP	SUBJECT	SPEARMAN RANK ORDER			PEARSON PRODUCT MOMENT		
			r	N	SIG	r	N	SIG
1978	1978	ENGLISH	.55	62	.001	–	–	–
		FRENCH (0120)	*	*	*	–	–	–
		PHYSICS	.76	44	.001	–	–	–
		MATHS	.59	56	.001	–	–	–
1978	1977	ENGLISH	.43	92	.001	.51	92	.001
		FRENCH (0120)	*	*	*	–	–	–
		PHYSICS	.49	51	.001	.52	51	.001
		MATHS	.37	81	.001	.38	81	.001
1977 + 1978	1977 + 1987	ENGLISH	.57	154	.001	.62	154	.001
		FRENCH (0120)	.70	73	.001	.73	73	.001
		PHYSICS	.60	109	.001	.64	109	.001
		MATHS	.57	136	.001	.60	136	.001

(*The equation, based on only twenty-two cases, is seen to be unreliable from the low value of the percentage of explained variance (15%).)

It must be remembered that the prediction equation is a way of combining scores to achieve a predicted result which correlates as highly as possible with the actual result. The first variable in the equation is the best single predictor, but it does not follow that out of all the remaining variables, the others that are included will be the next best single predictors. Of two variables that are highly correlated with the examination mark and also highly correlated themselves for the particular group under consideration, only one of the variables will appear in the equation. The amount of extra variance in the examination mark that can be explained by the second variable will be negligable when the first one is in the equation. The complete set of variables included in the equation between them account for as much variance as possible in the examination mark.

If one looks at Table 7 for the French Predictor for 1977 and 1978, one sees two extremely different equations (at face value). The correlation between C and VR in 1978 was 0.63, the correlation between VR and LU was 0.72 in 1977 and so it is no surprise to find LU as the main predictor in the combined equation (Table 7).

It is interesting to note that no creativity scores feature in the final equations and that, in general, the DAT scores account for most variance.

The standard error is an important measure, being the standard deviation of the difference between predicted and actual examination marks. The fact that in no subjects do we get (nor indeed do we expect to get) 100% variance would indicate a necessity for extreme caution in using knowledge of scores even from this extensive battery of tests to completely judge the potential of pupils. The reader will be aware of the caution that we have found necessary in our work with the equations at the source of a background criterion against which to assess teacher nomination of able pupils, but it is worth re-emphasising here. The equations could be used as a source of information to teachers in the identification process. If they are used in conjunction with other information - information about the pupils that is not entered into the equation - the predicted scores can be a useful piece of evidence to ability, but it would be extremely unwise to rank order pupils on their predicted scores and take no account of test error. The one or two marks that separate some pupils on this rank ordered list can be extremely misleading.

Another interesting point is that the main predictors for the subjects are quite different from one another. This reinforces the view that identification of the more able in a subject specific way is correct. None of the final equations are dominated by VN, which together give the general scholastic measure, correlating highly with IQ.

For completeness we list below the prediction equations for boys and girls separately (from the combined '77, '78 group).

TABLE 9

PREDICTION EQUATIONS FOR BOYS AND GIRLS
SEPARATELY (COMBINED GROUP)

BOYS

SUBJECT	EQUATION FOR CALCULATING PREDICTED	NO. OF CASES	% VARIANCE	STANDARD ERROR
ENG	1.064VR + 0.43SP + 40.48	78	35	14.55
FR (0120)	3.37LU - 8.42A - 0.84SR	29	42	20.67
PHYS	1.47MR + 2.27NA - 1.22CSA + 45.49	75	37	24.4
MATHS	2.89NA + 1.81AR - 7.29I + 0.97MR - 20.07	79	36	29.56

GIRLS

SUBJECT	EQUATION FOR CALCULATING PREDICTED	NO. OF CASES	% VARIANCE	STANDARD ERROR
ENG	1.34LU + 5.23G + 1.025NA + 0.38SP + 3.88	67	41	15.45
FR (0120)	1.88LU + 6.79E + 6.66C + 0.54SP - 90.77	38	58	17.09
PHYS	3.64MR + 25.93D + 1.94NA - 14.02E - 0.7SP - 180.85	16	87	9.3
MATHS	3.12NA + 1.46LU 6.47	50	22	32.45

The difference in the main predictor for English can be accounted for by the fact that for this group of boys the correlation between VR and LU is .52. Hence in all equations the main predictors are similar to those for the combined group.

From the results so far reported it was clear that the regression technique could be used with confidence to establish the criteria with which teachers' judgements can be compared. There is further evidence for the validity of this technique from the work of Guilford, Hoepfrer and Peterson (1965) who used various combinations of tests, notable among which was the DAT battery. The researchers established high correlations between prediction equations and scores in a variety of mathematics courses. Similar success for the method is reported by Cunningham et al (1978) using S.O.I. ability scores in regression analysis when resulting predictions were compared with assessments by teachers who were trained in the identification of able pupils.

Further evidence for the validity of the equations in Mathematics and French were gained from two sources. In Mathematics it was possible to check the validity in one school, where the scores were predicted at third form level for a group of pupils who entered O-level in the fourth form. The following table shows the rank ordered predicted scores of the puils and the actual grades achieved.

TABLE 10

PREDICTED SCORES AND ACTUAL O-LEVEL RESULTS FOR 30 PUPILS

PREDICTED SCORE	GRADE ATTAINED
203.81	A
185.48	A
185.14	A
177.15	A
168.20	B
165.94	B
165.83	A
162.82	A
157.52	A
150.38	B
150.20	B
148.75	B
148.71	B
148.05	B
147.77	D
146.67	B
146.12	B
145.95	A
144.79	A
144.77	B
142.42	B
141.06	B
141.02	A
138.68	B
136.56	B
135.61	B
130.43	B
128.83	[illegible]
124.37	C
121.68	E

The marks for these candidates were released in confidence by the Oxford Delegacy for Local Examinations and the Pearson Product Moment correlation with predicted score was .636 with an average error of 18.7. This should be compared with the percentage variance of 33% (approximate correlation of $\sqrt{.33}$= .575 and standard error 30.32) that is forecast from the prediction equations. The agreement between actual scores and predicted scores is thus seen to be considerably accurate, within the limits of the predicted standard error.

In French we were able to establish the Pearson Product Moment Correlation between predicted O-level score and a French Aptitude test[10]. The correlation was approximately 0.6 for a group of 130 pupils. We do not put this forward as any more than a small amount of extra evidence and further work could be done of this kind.

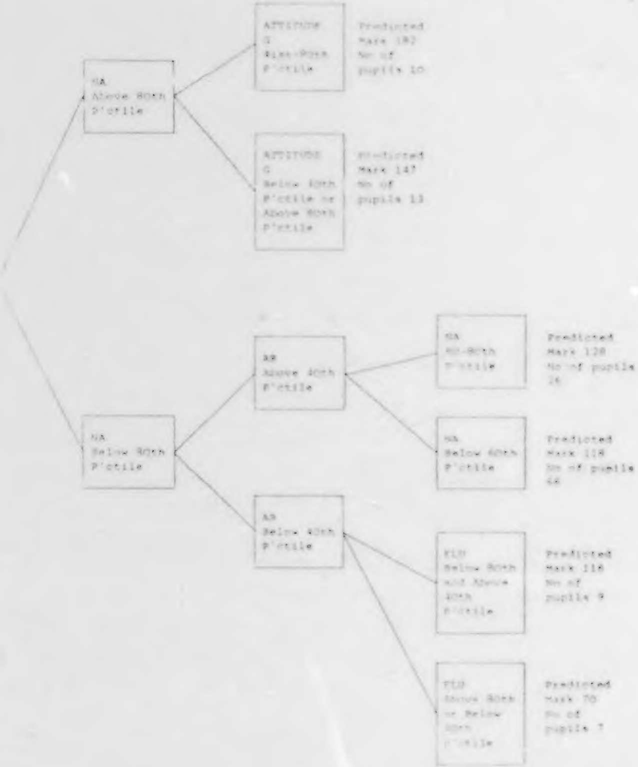
Finally, it is appropriate to point out that the prediction equation method was adopted after due consideration of a variety of other possible techniques.

First we must compare the method with a choice from the number of subject specific tests available. The prediction equation method allowed a large variety of aptitude and attitude scores to be incorporated into the work. They were entered into the data bank from which the equations were derived and so were 'available' for inclusion. Also they became a source to enable further study, particularly of the characteristics of the pupils selected by teachers, to be carried out. We were also able, by this method to test pupils in one way for all subjects. If the percentage of variance explained by the equations had not been so high, these advantages could not have been exploited, but our confidence in the equations, enabled us to exploit the advantages with confidence.

Second we investigated other statistical techniques with some energy. In particular two methods were looked at closely. Discriminant Analysis was investigated but was found to be most useful in the form that amounted to multiple regression analysis[11]. Segmentation Analysis[12] was also attempted. For this, the pupils are split first into two groups with respect to their scores on the most appropriate background test, (for example those scoring above and below the 80th percentile on NA). This is done in such a way that the variation in O-level score within the two groups is small, but between the two groups is as large as possible. These two groups are then split, each on a second background test score in a similar way. Further splits take place until no more meaningful splits can be made, so that in the end a variety of groups of pupils exist, each characterised by the manner of the splits, so that within the groups there is a small variation in O-level score. One can then use the categories of background tests which characterise these pupils to predict an O-level score for other pupils (the predicted score is the mean O-level score of the group that generated the result) - For example, after three 'splits' in the attempt to obtain mathematics predictors we obtained the following categories of pupils:

TABLE 8

SEGMENTATION ANALYSIS SPLITS FOR MATHEMATICS (UP TO THIRD SPLIT)


NA
Above 80th
P'ctile
ATTITUDE
G
31st-90th
P'ctile
Predicted
Mark 182
No of
pupils 10
ATTITUDE
G
Below 30th
P'ctile or
Above 90th
P'ctile
Predicted
Mark 147
No of
pupils 13
NA
Below 80th
P'ctile
AB
Above 40th
P'ctile
NA
40-80th
P'ctile
Predicted
Mark 128
No of pupils
26
NA
Below 40th
P'ctile
Predicted
Mark 118
No of pupils
46
AB
Below 40th
P'ctile
FLU
Below 80th
and Above
30th
P'ctile
Predicted
Mark 116
No of
pupils 9
FLU
Above 80th
or Below
30th
P'ctile
Predicted
Mark 70
No of
pupils 7

The method may have had the potential of a very sound and useable strategy but we could not pursue it because it was found that very large numbers of pupils would be needed to make 'splits' beyond the first one meaningful. We include this brief description first to show that the prediction equation method emerged as the best that we could find, rather than the only one that was considered, and second to suggest that a further line of fruitful research might be to follow up a prediction method based on segmentation analysis.

Remaining notes for Appendix 1

Notes 1-5 are to be found at the foot of the fold out sheet.

6) Vernon, P.E., *Secondary School Selection*, London, Methuen, 1957.

7) For a full discussion of this scale, see Bennett, G.K., Seashore, H.G. and Wesman, A.G., *5th Manual for the Differential Aptitude Tests*. The Psychological Corporation, 1974.

8) Postlethwaite, K.C. and Denton, F.C.J., *Streams for the Future?*, Pubansco, Banbury, 1978.

9) See, for example, the work of Rosenthal R. and Jacobson L., *Pygmalion in the classroom*, Holt, Reinhard and Winston 1968. Their methodology has been widely criticised (Elashoff and Snow, 1971, *Pygmalion Reconsidered*; Ohio, C.A. Jones) but later studies summarised by Rogers (*The Social Psychology of Schooling*, London, R.K.P.) have confirmed that, though the self-fulfilling prophecy effect may not always happen, it certainly can happen.

10) This was developed by Miss Ann Miller for the OERG Modern Languages Project. Her test was a better predictor for end of year tests than the well known Pimsleur Language Aptitude Battery on the sample of about 50 pupils on whom it was applied (correlation of approx. 0.7 with end of year test score.)

11) The SPSS Manual describes how this can be so in an appropriate way.

12) See Sonquist J.A. and Morgan J.N. *The Detection of Interaction Effects* (1964) Mono No 35 Survey Research Centre. University of Michigan, for a fuller description of the algorithms.

APPENDIX 2

Forms used by teachers for listing top 10% of pupils

OERC MORE ABLE PUPIL PROJECT

Teacher based identification of the most able 10% of pupils in each of four subject areas

In accordance with the outline programme circulated to schools earlier we are now asking heads of departments of mathematics, English, physics and French to provide us with a list of names of the most able 10% of 3rd year pupils in their subject. We hope that the following notes will make clear exactly what we are seeking in making this request.

1) The list should consist of that 10% of 3rd year pupils whom you expect to have the greatest potential for further work in your subject. They may not necessarily be the 10% who come top in tests or examinations.

2) We are not seeking comparisons made on any national or county-wide basis. We are just interested in the most able 10% within your own school.

3) You may obviously make use of any data which you feel would be useful in completing this task.

If you need any further information please do not hesitate to ring us on Oxford 54121 Ext. 276 or 277.

Please give the names of this top 10% on the following pages.

For the purpose of enabling us to analyse with greater precision the way in which teacher-based identification works we would be grateful if you could attempt to subdivide this 10% target group.

Please give below the names of those pupils in the upper half of the main target group (i.e. the top 5% in your subject).

So that we can follow up details of this exercise please complete the following questions.

How was the main target group chosen?

nomination by teacher of top set [ ]

nomination by several teachers followed by production of list by head of dept. [ ]

discussion in departmental meeting [ ]

other (please specify) ________________ [ ]

____

____

What information was used?

teachers' own records [ ]

school examination results [ ]

school reports [ ]

objective tests [ ]

other (please specify) ________________ [ ]

____

____

Which teachers were involved?

____

____

____

____

____

We are very grateful for your help in completing these forms. We hope to be able to feed back the outcome of this survey in the second half of the Easter Term.

CLIFF DENTON
KEITH POSTLETHWAITE
Research Officers.

School ____________________

Subject Maths( )   French( )   English( )   Physics( ) (please tick)

Please give your top 10% group here - it need not be in rank order.

Please indicate with a * any pupil whom you regard as exceptional (anyone who, if they were to continue with the subject could be expected to make a serious contribution to the study of or application of that subject). (You may feel that no one is this category.) Please also indicate with a ? those pupils whom you feel to be on the borderline of this group.

If there are other pupils whose exclusion from this group caused you difficulty you could list them here:

APPENDIX 3

Further details concerning the [illegible] pupil sample

Characteristics of pupils in the sample

TABLE 1

DIFFERENTIAL APTITUDE SCORES (Year 1 of Study)

	SAMPLE			NFER STANDARDISATION		
	mean	sd	n	mean	sd	n
VR	22.1	8.89	2151	22.5	10.0	2394
NA	13.9	6.44	2157	[illegible].9	6.6	2396
VN	35.5	13.82	2148	35.4	15.2	2394
AR	30.9	9.82	2147	29.7	11.0	2402
CSA	42.5	11.97	2149	38.2	11.9	2339
MR	41.0	9.95	2158	39.5	9.8	2339
SR	29.3	10.58	2167	26.4	10.2	2379
SP	56.1	15.87	2162	55.5	16.2	2327
LU	26.5	9.28	2165	25.9	9.6	2325

TABLE 2

DISTRIBUTION IN THE SAMPLE
REGISTRAR GENERAL CLASSIFICATION

Class	1	2	3	4	5
Number	118	440	1108	270	73

We would not wish to claim, from the figures above, that our sample is truly representative of the country as a whole. However, it is clear from these figures that the sample is not as heavily biased towards high ability and towards the professional end of the social spectrum as might be imagined from a moment's reflection on the character of the county of Oxfordshire.

Indeed, the distribution of social class in School 02 was as follows:

Class	1	2	3	4	5
No. of pupils	2	22	136	53	15

This distribution, heavily weighted away from the high social classes, is significantly different from the distribution in the rest of the sample (χ^2 = 37.27 df = 4 $p \leq 0.001$) and indicates a school with substantial numbers of pupils from lower social classes.

TABLE 3

SCHOOL TYPES AND YEAR ONE INVOLVEMENT

Category	Schools	
11+ entry	00 01 03 04 06 07 08 13[1]	
13+ entry	02 05[2] 09	
Size within the 11+ group		
large	13[1]	(2000+ pupils)
medium	00 01 04 06 07	(1000 - 1500 pupils)
small	03 08	(less than 1000 pupils)
Control[3]	00 05 06 07 08 09	
Experimental (Year 1)	01 02 03 04 13[1]	

School involvement in year two

TABLE 4

MAIN ACTIVITY[4] THAT EACH SCHOOL WAS INVOLVED IN FROM THE THREE TIERS[4] OF FIELDWORK IN THE SECOND YEAR

SCHOOL	ENGLISH	FRENCH	PHYSICS	MATHS
00	3	3	3	2 & 3
01	1	1	1	1
02	1	1	1	1
03	2	2	2 & 3	3
04	3	3	2	2
05	2	2	2	3
06	1	1	1	1
07	1	2	2	3
08	3	3	2 & 3	3
09	3	3	2 & 3	2
10	2	1	2	3
11	2	1	2	1
12	2	1	2	2

Notes for Appendix 3

1) 'Schools' 10, 11 and 12 are component parts of one large school (13).

2) School 02 is a single sex school (boys). School 05 is a single sex school (girls).

3) The division of schools into checklist and non-checklist groups was done as follows. Of the three 13-18 schools one of the two single sex schools was randomly chosen for the checklist group. Two of the four 11-18 schools with more than one thousand pupils were randomly chosen for the checklist group and two of the three smaller 11-18 schools were randomly chosen to use checklists. In this way, as far as was possible, in view of the nature of the schools working with us, examples of each type of school were allocated to the checklist and non checklist groups. The one 13-18 mixed school and the one 11-16 school were placed in the non checklist group in the 1st year. This allowed the possibility of a further comparison in 1981/2, the 2nd year of research, when these schools could be asked to use checklists for the first time (they were, therefore, being used as their own controls).

4) Most schools co-operated in tier 1 for the majority of subjects.

5) Tier 1: retrospective case study involving interviews to discover reasons for mismatch between test and teacher assessment in year 1.
   Tier 2: Structured work in year 2 to discover characteristics of pupils which teachers use in their identification strategy.
   Tier 3: Checklist work.

APPENDIX 4

Further statistical tables

We have included, in this appendix, the results of analyses that were carried out alongside those reported in the main text. Often alternative approaches were available, particularly with regard to tests of significance. We have given the most appropriate results in the main text and committed these other results to this appendix.

TABLE 1

Table Showing Those Variables Which Yielded Significant Correlation With Teacher Selection (Round 1) When Compared With Rest of Year Group

SCHOOL	SUBJECT	NO OF CASES	VARIABLE	PEARSON CORRELATION COEFFICIENT	SIGNIFICANCE
00	ENGLISH	157	A	0.2735	***
			B	0.1740	*
			C	0.3000	***
			G	0.1968	**
			VR	0.4466	***
			NA	0.3977	***
			VN	0.4664	***
			AR	0.3076	***
			CSA	0.1662	*
			MR	0.1923	**
			SR	0.3466	***
			SP	0.4133	***
			LP	0.4816	***
00	FRENCH	157	B	0.1889	*
			C	0.3534	***
			G	0.1385	*
			VR	0.5615	***
			NA	0.5826	***
			VN	0.6256	***
			AR	0.3821	***
			MR	0.3338	***
			SR	0.4081	***
			SF	0.4313	***
			LP	0.5345	***
00	PHYSICS	157	A	0.2188	**
			B	0.1928	**
			C	0.4139	***
			D	0.1612	*
			G	0.1649	*
			VR	0.4709	***
			NA	0.4837	***
			VN	0.5221	***
			AR	0.3689	***
			MR	0.3284	***
			SR	0.4946	***
			SF	0.3352	***
			LP	0.4409	***
00	MATHS	157	FLU	0.1581	*
			FLEX	0.1976	**
			ORIG	0.1800	*
			C	0.3868	***
			VR	0.4772	***
			NA	0.5858	***
			VN	0.5745	***
			AR	0.3563	***
			CSA	0.2023	**
			MR	0.3083	***
			SR	0.3855	***
			SF	0.3874	***
			LP	0.4376	***

* significant at 5% level ** significant at 1% level *** significant at 0.1% level

TABLE 2

TABLE SHOWING THOSE VARIABLES WHICH YIELDED A SIGNIFICANT VALUE OF CHI-SQUARED. TEACHER SELECTION VERSUS REST OF YEAR 1 (ROUND 1) SCHOOL 00

	SEX	NEATNESS	SOCIAL CLASS
ENGLISH	nsd	nsd	7.23 *
FRENCH	nsd	nsd	14.07 **
PHYSICS	nsd	nsd	10.95 **
MATHS	nsd	nsd	7.73 **

The above tables, for one school only, show a large number of significances. It would be unhelpful to show the results for all schools where in each case a large number of variables were associated with teacher selection. A scrutiny of the correlation coefficients indicates that some variables have much stronger correlation than others. Indeed an examination of the intercorrelations between the variables for a particular school further suggests that some variables appear as factors in the selection process only because of their links with other variables. The technique of partial correlation allows us to 'control' for one variable and then measure the new correlation between teacher selection and the other variables. The technique, therefore, allows us to discover the variables that are strongest in the characterisation of the teacher selected group of pupils. The variables that remain after partial correlation are shown below. The appropriate Pearson correlation coefficients are also quoted.

TABLE 3

TABLE SHOWING THE PEARSON CORRELATION COEFFICIENT FOR THOSE VARIABLES THAT MOST STRONGLY DISTINGUISH BETWEEN THE TEACHER SELECTED GROUPS AND THE WHOLE YEAR GROUP (ROUND 1)[1]

SCHOOL	SUBJECT	SIGNIFICANTLY CORRELATED VARIABLES
00 (157)	ENGLISH	VN (0.45) NA (0.40) SR (0.34) LU (0.48)
	FRENCH	VN (0.63)
	PHYSICS	C (0.41) VN (0.52) SR (0.49)
	MATHS	NA (0.54) SR (0.38)
01 (218)	ENGLISH	VN (0.54) LU (0.53) ORIG (0.47)
	FRENCH	VN (0.55) NA (0.53)
	PHYSICS	NA (0.40) VN (0.40) MR (0.36)
	MATHS	VN (0.53)
02 (192)	ENGLISH	SF (0.47) LU (0.53)
	FRENCH	LU (0.44) SF (0.43) VN (0.44) C (0.30) D (0.34)
	PHYSICS	VN (0.47) C (0.29) A (0.23)
	MATHS	NA (0.56) D (0.33)
03 (162)	ENGLISH	FLEX (0.40) VN (0.33) SF (0.39) B (0.28)
	FRENCH	LU (0.44) NA (0.45) CSA (0.37)
	PHYSICS	VN (0.44) MR (0.38)
	MATHS	NA (0.53) A (0.33)
04 (196)	ENGLISH	LU (0.58) SF (0.53) B (0.33)
	FRENCH	LU (0.59) A (0.57) D (0.32)
	PHYSICS	SR (0.28) A (0.27) LU (0.35)
	MATHS	NA (0.54) C (0.33)
05 (166)	ENGLISH	VN (0.42) C (0.35) SF (0.43)
	FRENCH	VN (0.50) SF (0.43) LU (0.48)
	PHYSICS	VN (0.49) FLEX (0.30) C (0.30) MR (0.44)
	MATHS	NA (0.64) C (0.39)
06 (202)	ENGLISH	LU (0.54) B (0.44) C (0.41) FLEX (0.38) NA (0.40)
	FRENCH	LU (0.53) D (0.37)
	PHYSICS	VN (0.52) B (0.44)
	MATHS	NA (0.52) C (0.42)
*07 (384)	ENGLISH	VN (0.48) C (0.58) SF (0.41) LU (0.43)
	FRENCH	VN (0.39) LU (0.40) SF (0.34) A (0.26) C (0.35)
	PHYSICS	VN (0.45) SR (0.45)
	MATHS	VN (0.57) C (0.43) SR (0.44)
08 (130)	ENGLISH	VN (0.45) SF (0.49)
	FRENCH	LU (0.54) VN (0.56)
	PHYSICS	VN (0.35) FLEX (0.33) ORIG (0.32) B (0.27)
	MATHS	NA (0.54)
09 (290)	ENGLISH	VN (0.55) LU (0.58) ORIG (0.30)
	FRENCH	ORIG (0.43) VR (0.55) SF (0.50) LU (0.57)
	PHYSICS	VN (0.56) LU (0.55) MR (0.43) C (0.27)
	MATHS	VN (0.54) LU (0.53)
10 (128)	ENGLISH	VN (0.50) LU (0.60)
	FRENCH	LU (0.56)
	PHYSICS	VN (0.48) B (0.50) C (0.40)
	MATHS	VN (0.50)
11 (327)	ENGLISH	VN (0.50) B (0.29)
	FRENCH	VN (0.56)
	PHYSICS	VN (0.41) SR (0.42)
	MATHS	VN (0.52) NA (0.52) VR (0.42) SR (0.43)
12 (329)	ENGLISH	LU (0.46) SF (0.43) NA (0.35)
	FRENCH	SF (0.37)
	PHYSICS	VN (0.52)
	MATHS	VN (0.56)

---

[1] The significance of correlation in all cases is at the 0.1% (***) level. [illegible]-group correlations are in brackets next to the variable name. The number of pupils is in brackets under the school code.

A second way to assess the characteristics of the selected group is to make a comparison of mean scores on each of the background tests of the selected group against the non-selected group. A significant difference as indicated by the t-test indicates a strength (or weakness) on that characteristic for the selected group. The pattern of significantly different mean scores reflects the pattern of significantly different correlations and just serves to confirm the characteristics of the teacher identified groups. Again, we see that this group is not superior on 'everything'. Higher mean scores on DAT is the main characteristic and the different additional characteristics are associasted with different schools. Attitudes B (Relationship with teacher) and C (Academic Self Image) are again prominent, however, in nearly all of the schools for all subjects. Table 4 summarises the results across the schools.

TABLE 5

TOTAL NO OF SCHOOLS SHOWING SIGNIFICANTLY DIFFERENT F OR T VALUES WHEN THE TEACHER IDENTIFIED GROUP IS COMPARED WITH THE REST OF THE YEAR GROUP[1]

	ENGLISH			FRENCH			PHYSICS			MATHS		
	NO OF SIG. ON			NO OF SIG. ON			NO OF SIG. ON			NO OF SIG. ON		
	F	t	F OR t	F	t	F OR t	F	t	F OR t	F	t	F OR t
CREATIVITY												
FLU	2	5	6	0	4	4	0	5	5	1	4	5
FLEX	2	8	8	1	3	3	1	6	6	0	2	2
ORIG	1	7	7	0	2	2	1	5	5	1	4	5
ATTITUDE												
A	1	8	9	1	10	11	2	7	8	2	6	7
B	0	12	12	0	11	11	0	10	10	0	10	10
C	0	12	12	2	10	12	2	10	11	1	11	11
D	1	1	1	5	5	8	3	4	5	1	3	3
E	0	5	5	1	5	6	2	3	3	1	2	2
F	0	2	2	1	1	2	0	3	3	1	2	3
G	5	6	9	5	8	8	1	5	5	3	6	6
I	2	5	6	0	6	6	0	3	3	0	3	3
DAT												
VR	1	15	15	1	15	15	0	15	15	0	15	15
NA	2	15	15	0	15	15	0	15	15	0	15	15
VN	1	15	15	0	15	15	1	15	15	0	15	15
SR	10	15	15	7	15	15	10	15	15	10	15	15
CSA	3	11	12	3	10	10	2	5	7	5	10	14
MR	0	11	11	1	10	10	0	15	15	0	15	15
SE	1	11	11	2	12	12	0	15	15	0	15	15
SP	6	15	15	5	15	15	0	15	15	1	15	15
LU	5	15	15	3	15	15	1	15	15	0	15	15
NEATNESS (NT)	1	9	9	0	9	9	0	3	3	0	4	4
SOCIAL CLASS (SC)	0	3	3	1	10	11	0	3	3	0	2	2

---

[1] The normal statistical procedure is to check whether the two groups have a different variance (spread of scores) by the F test, then to check if there is a significant difference in the mean score by the t test. A non-significant F value is required in order to carry out a valid t test, though a significant F or t indicates that one or other difference is present.

169

The only comment, in addition to the fact that Table 4 supports the study of characteristics by Correlation Analysis is to draw attention to the large number of significant values of attitude F and of AR (Abstract reasoning) in all subjects. The significance was in most cases beyond the 1% level and in every case indicated a smaller spread of scores for the identified group. An interpretation of this would seem important though seem important though there is no obvious explanation. It may have been that there was something that is observable in the character of the children which correlated well with AR, and that our tests did not reveal.

Another enlightening exercise is that of Multiple Regression. If we give pupils a score of 1 if they are on the teacher's subject list as in the top 10% and 0 otherwise, we can then apply the Multiple Regression technique to derive a prediction formula for selection by teachers. This again supports the correlation technique but gives additional information in the form of percentage variance explained by the prediction equation which contains the variables that (weighted according to the regression formulae) best characterise the selection process.

REGRESSION FORMULAE WHICH PREDICT FOR TEACHER SELECTION FROM WHOLE SCHOOL GROUP (NO. OF PUPILS IN BRACKETS) ROUND 1

SCHOOL	SUBJECT	EQUATION	% AGE VARIANCE	STANDARD ERROR
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]
[illegible]	ENGLISH	[illegible]	[illegible]	[illegible]
	FRENCH	[illegible]	[illegible]	[illegible]
	PHYSICS	[illegible]	[illegible]	[illegible]
	MATHS	[illegible]	[illegible]	[illegible]

Again, as in Chapter Four when looking at Partial Correlation results, on the whole these are 'reasonable' characteristics, though the strength of VN compared with the more appropriate measure (LU for English, LU for French, MR for physics and NA for mathematics) is apparent and indicates a tendency to choose pupils with general academic ability. School 03 (Flex for English) and Schools 05 and 08 (Spelling for English), are a little more extreme. physics, again looks to be most strongly biased in selection strategy to VN.

We now turn to the characteristics of the teacher selected pupils which distinguished them from the rest of the B Group pupils. The correlation between teacher selection and the significantly correlated background test scores for B Group pupils only are shown in Table 7.

The percentage variance explained in each case is encouraging in that we have found variables that characterise the selected group quite strongly. To understand the equations in full is not easy and to suppose that the selection process was equivalent to forming a regression equation would be hard to justify and a matter for careful psychological study. There is again support, however, for the characteristics of the identified group that were apparent from partial correlation, as one can verify if one compares the equations with those in Chapter Four. The standard error in all cases is approximately .25, which again gives one some confidence in the method.

It is useful [illegible]ok at the variable that correlates highest with teacher selection and at the percentage variance explained by it[1].

TABLE 6

PERCENTAGE VARIANCE EXPLAINED BY THE 'STRONGEST'
VARIABLE THAT CHARACTERISES THE SELECTED 10% GROUP

SCHOOL	ENGLISH	FRENCH	PHYSICS	MATHS
00	LU (23%)	VN (39%)	VN (32%)	NA (32%)
01	VN (27%)	VN (37%)	VN (14%)	VN (30%)
02	LU (27%)	LU (21%)	VN (21%)	NA (30%)
03	FLEX(16%)	NA (21%)	VN (16%)	NA (29%)
04	LU (27%)	LU (16%)	SR (10%)	NA (27%)
05	SP (19%)	VS (27%)	VN (24%)	NA (45%)
06	LU (30%)	LU (30%)	VN (27%)	NA (27%)
07	VN (23%)	LU (15%)	NN (22%)	VN (30%)
08	SP (25%)	VR (13%)	SR (14%)	NA (25%)
09	LU (34%)	LU (33%)	VN (32%)	VN (30%)
10	LU (41%)	LU (34%)	VN (22%)	VN (28%)
11	VN (25%)	VN (30%)	SR (18%)	NA (25%)
12	LU (34%)	SP (17%)	VN (34%)	VN (35%)

TABLE [illegible]

TABLE SHOWING THE PEARSON CORR. COEFF. FOR THOSE VARIABLES
THAT MOST STRONGLY CHARACTERIZE TEACHER SELECTION IN THE 'B' GROUP (ROUND 1)

SCHOOL	SUB (NO OF PUPILS)	SIGNIFICANTLY CORRELATED VARIABLES[1]
00	Eng (25)	CSA (0.3963)
	Fr (20)	SW (0.6313) VW (0.6752)
	Phys (14)	CSA (0.5871) T (0.4723)
	Maths (29)	AP (0.3706*)
01	Eng (25)	NONE
	Fr (21)	CSA (0.5667) A (0.3771)
	Phys (46)	NA (0.3416**) SW (0.2882*)
	Maths (15)	CSA (0.4359**) A (0.4151**)
02	Eng (42)	R (-0.3714**) SW (0.3563**) IU (0.3319*)
	Fr (13)	NONE
	Phys (30)	V (0.6545) NA (0.4333**)
	Maths (30)	T (0.4635**) NA (0.5276)
03	Eng (30)	FLU (0.5344**) FLEX (0.4835**) ORIG (0.4985**)
	Fr (19)	A (0.6532**)
	Phys (33)	VW (0.5424) SP (0.5732) IU (0.4408**)
	Maths (23)	A (0.5545) NA (0.4254*)
04	Eng (38)	IU (0.3127**)
	Fr (20)	A (0.5573*)
	Phys (43)	CSA (0.3132*) VW (0.3305*)
	Maths (34)	R (0.5262) NA (0.4378**)
05	Eng (27)	NONE
	Fr (13)	T (-0.6565) IU (0.4833*)
	Phys (22)	SE (0.5272*) VW (0.5509**)
	Maths (15)	NA (0.6273**) VW (0.4948*)
06	Eng (33)	IU (0.4565**) R (0.3178*) T (0.3380*) D (0.3428*)
	Fr (23)	IU (0.5652**) A (0.5518**) R (0.5610**)
	Phys (41)	R (0.3131) VW (0.5608)
	Maths (30)	T (0.3769**) SW (0.3630*)
07	Eng (38)	FLEX (0.2830*) SW (0.3182*)
	Fr (24)	VW (0.4535*)
	Phys (36)	SW (0.5274) F (0.5030) I (0.3455*)
	Maths (38)	SE (0.5048)
08	Eng (20)	SP (0.8024*)
	Fr (18)	NONE
	Phys (24)	FLEX (0.5315**) SE (0.3590*)
	Maths (23)	CSA (0.5968) A (0.5456**) R (0.5533**)
09	Eng (38)	IU (0.3969**) CSA (-0.3691*) ORIG (0.3720*) S (-0.3759*)
	Fr (28)	CSA (-0.5139*) SP (0.4837**)
	Phys (53)	VW (0.4630) C (0.3756**)
	Maths (36)	AP (0.4013**) IU (0.3691**)
10	Eng (20)	VW (0.5323**) AP (0.6081**) NA (0.5231**)
	Fr (13)	A (0.5120*)
	Phys (22)	SW (0.5391**) AP (0.5302**) IU (0.4336)
	Maths (19)	SW (0.4712*) AP (0.5258**)
11	Eng (19)	ORIG (0.4673**) FLEX (0.4780*)
	Fr (14)	C (0.7464) CSA (0.6783**)
	Phys (35)	SW (0.6075) CSA (0.5282) T (0.4551**)
	Maths (21)	NA (0.5257**) VW (0.6135) AR (0.6054**) CSA (0.5476**)
12	Eng (36)	NONE
	Fr (20)	SP (0.5702**)
	Phys (17)	IU (0.7379)
	Maths (20)	AR (0.4369*)

---

[1] All significances are at 0.1% (***) level unless shown otherwise.

The characteristics of the teacher selected group within the 'B' group are again supported by the difference in means of the selected group and the rest from test calculations the results are shown in Table 8. The variation in characteristics from school to school that we observed from the correlation analysis in Chapter Four is shown up in the table.

The small numbers involved would lead us to treat our t and r comparisons with caution, but they are included to add weight to our previous findings. The technique of devising multiple regression equations on the B group alone would be subject to similar small number cautions but it is useful to make up a table of the percentage variance explained by the variable that enters the equation first (Table 9).

[illegible]

TABLE 9

PERCENTAGE VARIANCE AND STANDARD ERROR OF PREDICTION EXPLAINED BY THE 'STRONGEST' VARIABLE THAT CHARACTERISES SELECTION FROM THE 'B GROUP' (ROUND 1, YEAR 1)

School	Eng Vble (%Var) SE	Fr Vble (%Var) SE	Phys Vble (%Var) SE	Maths Vble (%Var) SE
00	CSA (35%) 0.40	VN (44%) 0.37	CSA (39%) 0.32	AR (11%) 0.40
01	BRX (14%) 0.45	A (31%) 0.34	NA (10%) 0.43	CSA (17%) 0.44
02	SP (12%) 0.43	E (9%) 0.45	B (27%) 0.40	NA (24%) 0.43
03	FLU (26%) 0.39	A (35%) 0.36	SP (27%) 0.38	A (29%) 0.36
04	LU (14%) 0.43	A (21%) 0.33	CSA (8%) 0.36	E (21%)0.39
05	NONE[1]	E (40%) 0.37	SP (29%) 0.37	NA (18%) 0.38
06	LU (18%) 0.41	B (30%) 0.42	VN (30%) 0.41	C (15%) 0.44
07	FLEX (10%) 0.36	VN (21%) 0.41	SR (26%) 0.29	SR (24%) 0.43
08	SP (14%) 0.44	NONE	FLEX (24%) 0.42	CSA (33%)0.41
09	C (14%) 0.46	CSA (23%) 0.43	VN (18%) 0.46	AR (11%) 0.46
10	AR (33%) 0.40	D (47%) 0.36	LU (37%) 0.38	VN (32%) 0.42
11	ORIG (13%) 0.37	C (52%) 0.36	SR (35%) 0.33	VN (34%) 0.38
12	NT (14%) 0.46	SP (33%) 0.31	LU (49%) 0.36	AR (19%) 0.46

---

[1] The Equation could not be formed in this case because of no valid predictor among the tests.

The first observation is that the standard error of regression here is much larger than for the Multiple Regression Method applied to the whole school group, confirming the conclusion that selection from within the 'B' group is difficult to predict and reflecting something of an indeterminate strategy[2].

The French predictions confirm the view that 'good' attitudes (in this case not specifically to the subject) characterise the selected groups quite strongly (9 out of 13 schools). It is difficult to make further comments on other strategies that were not made in Chapter Four except to repeat that different school/teacher selections are biased in their own peculiar ways and that there is a clear need for further evidence from within schools to supplement this background work.

Correlation between teacher nomination and predicted O-level score

In Chapter Three we reported on a method of taking account of test error. This helped to reduce the difficulties which always exist when analysis is based on dividing up an ordered list. One difficulty is, of course, that very few marks often separate people near the dividing line and therefore, in our context, false distinctions are made between nominating a pupil with a score of 120 who is above the line and a pupil with a score of 119 who happens to be below. Another difficulty is that all nominations 'above the line' are treated in the same way so that, in the work reported above, a teacher nominating the top 6 pupils out of a group of 12 above the line, would be judged to be equally accurate as a teacher nominating the lowest 6.

In an attempt to overcome these problems we have supplemented the analyses reported in Chapter Three by the analysis reported here which is based on the correlation between teacher judgement and predicted O-level score. No 'line' is referred to in this analysis which may therefore give a finer view of test/teacher agreement.

ROUND 1 (Year 1)

In Table 10 we give the values of the point-biserial correlation coefficients between teacher judgement and test-predicted O-level score. All pupils in the schools were included in this analysis.

TABLE 10

CORRELATIONS BETWEEN TEACHER JUDGEMENT AND TEST PREDICTED O LEVEL SCORE FOR ALL PUPILS IN EACH SCHOOL
(Year 1, Round 1)

SCHOOL	ENGLISH	FRENCH	PHYSICS	MATHS
00	.57 (145)	.49 (140)	.43 (149)	.52 (143)
01	.54 (205)	.44 ([illegible]00)	.43 (205)	.51 (201)
02	.51 (170)	.43 (167)	.42 (173)	.52 (171)
03	.50 (138)	.39 (135)	.38 (135)	.42 (136)
04	.55 (174)	.29 (174)	.14 (173)*	.47 (172)
05	.46 (144)	.44 (143)	.46 (146)	.59 (147)
06	.58 (196)	.42 (196)	.43 (197)	.48 (197)
07	.49 (220)	.37 (222)	.28 (222)	.50 (218)
08	.50 (124)	.34 (122)	.30 (121)	.48 (122)
09	.60 (189)	.51 (190)	.51 (189)	.53 (194)
10	.60 (118)	.50 (118)	.47 (119)	.47 (116)
11	.49 (112)	.44 (110)	.31 (111)	.58 (111)
12	.46 (116)	.22 (116)**	.30 (118)	.51 (118)

* = sig at 5% level

** = sig at 1% level

All other correlations were significant at the 0.1% level

All correlation coefficients in this table were statistically significant. Except where indicated this was at the 0.1% level. These significances in all schools in all subjects, provided evidence for a definite link between test and teacher judgements.

Since each school's third year can be treated as a random cluster sample from the same population (third year pupils in Oxfordshire) it is reasonable to find average correlation coefficients from the full set of school coefficients quoted in Table 10. This has been done by transforming the correlation coefficients into Z scores, finding the mean Z and transforming back[3]. In Table 11 average coefficients for control and experimental schools are also reported separately.

TABLE 11

MEAN CORRELATION COEFFICIENTS

	ENGLISH	FRENCH	PHYSICS	MATHS
All Schools	.52	.41	.39	.50
Control Schools	.53	.41	.39	.50
Experimental Schools	.51	.40	.39	.50

Several points emerge from these tables in support of findings reported elsewhere. First the larger coefficients in mathematics and English indicated greater agreement between tests and teachers in these subjects than in Physics and small differences between schools except in Physics. It is interesting to note that, in Physics, in agreement with the result reported in Chapter Three, School 04 showed least agreement between test and teacher, and School 09 showed greatest agreement.

Finally the similarity of coefficients in the control and experimental schools confined the lack of difference between these schools that was reported in Chapter Five.

ROUND 2 (Year 1)

As a supplementary check on the overall match between teacher and test judgments in Round 2 we repeated the study of the correlation between teacher nomination and predicted O-level score on Round 2 data. Table 12 indicates the correlation coefficients for each school and Table 13 the average coefficient.

TABLE 12

CORRELATION COEFFICIENTS BETWEEN ROUND 2 TEACHER JUDGEMENT AND TEST PREDICTED O LEVEL SCORE FOR ALL PUPILS IN EACH SCHOOL

SCHOOL	ENGLISH	FRENCH	PHYSICS	MATHS
00	.56 (142)	.49 (137)	.46 (146)	.57 (140)
01	.57 (265)	.35 (200)	.48 (205)	.49 (201)
02	-	.37 (164)	.42 (170)	.54 (168)
03	.51 (133)	.36 (130)	.25 (130)	.41 (132)
04	.49 (174)	.40 (174)	.34 (173)	.52 (172)
05	.45 (143)	.45 (142)	.42 (145)	.64 (146)
06	.61 (195)	.40 (195)	.46 (196)	.49 (196)
07	.49 (217)	.40 (219)	.37 (219)	.54 (215)
08	.47 (124)	.45 (122)	.42 (121)	.57 (122)
09	.62 (186)	.52 (187)	.36 (186)	.54 (190)
10	.60 (115)	.50 (115)	.43 (116)	.52 (113)
11	.50 (110)	-	.44 (109)	.53 (109)
12	.43 (112)	.35 (112)**	.30 (114)	.48 (114)

TABLE 13

MEAN CORRELATION COEFFICIENTS

	ENGLISH	FRENCH	PHYSICS	MATHS
All Schools	.53	.42	.40	.52
Control Schools	.55	.45	.41	.53
Experimental Schools	.51	.40	.39	.52

We can compare Tables 12 and 13 with Tables 10 and 11. There is no evidence here to suggest that checklists have resulted in any improvement in the match between teacher and test judgement and adds to evidence of a similar kind in Chapter Five.

Relationship between those pupils identified by teachers and appropriate test scales (Year 1, Round 1)

We show here the percentage of pupils who were in the top 10% on a particular test scale who were selected by teachers.

Table 14 shows these figures for the VN scale and for the best single predictor (the test most highly correlated with O-level score). All figures are for Round 1, Year 1.

TABLE 14

	ENGLISH	FRENCH	PHYSICS	MATHS
% of top 10% on VN also identified by teachers	45	50	47	58
% of top 10% on best single predictor also identified by teachers	57	50	38	59
Best single predictor	LU	LU	MR	NA

The idea that undue emphasis might be placed by teachers on VN, was reported in Chapter Four and is supported by the Physics results in this table. This table is too crude to indicate the effect in other subjects where the correlation between VN and the best single predictor is high (above 0.80).

Notes for Appendix 4

1) Since percentage variance is approximately the square of the correlation coefficient, there is little information here that could not be got from Table 3, except that we have identified the strongest correlation.

2) Recall that prediction is based on a score of 1 for identification and 0 otherwise. The standard error is about 0.4 in most cases.

3) This procedure is explained in Guilford, J. and Fruchter B. (1978).

APPENDIX 5

Instructions to teachers and checklists that were used in years 1 and 2 that have not been included elsewhere in the report

OERC MORE ABLE PUPILS PROJECT

Second round of teacher based identification of the most able ten percent of pupils in each of four subject areas (Year 1)

1. Introduction

1.1 The first two stages of the research programme (test of pupils and the first round of teacher based identification of the top 10% of pupils) are now largely complete and we are very grateful for the help you have given us in bringing the research to this stage. We now need to ask for your help in carrying out the last major school-based task. This is the second round of teacher based identification and is a crucial part of the project. From the results we hope to develop some tools which will be of practical value in guiding identification in schools.

1.2 As in the first round of identification we are interested in listing that ten percent of 3rd year pupils who show the greatest potential for future achievement in your subject. They may not necessarily be the 10% who come top in third year tests or examinations. Again, as last time, we are not seeking any comparisons made on a national or county-wide basis. We are just interested in the most able 10% within your own school.

1.3 The key difference between the first and second rounds of identification is that in this second round we are asking you to make use of a checklist which is intended to provide an additional source of information about the pupils. This is the first time that such checklists have been used in this way so the items on the list, and the detailed way in which it is interpreted, can only be tentative at this stage. We hope that you will feel free to comment on any aspect of the lists so that better versions can emerge. We would, however, ask you to keep as closely as you can to the instructions given while you complete the task this time.

2. General instructions for the use of the Checklist

2.1 Target group

Please make a list of roughly the top 20 per cent of pupils according to the criteria explained in 1.2 above. Please feel free to vary the figures of 20% if this will allow you to work with convenient groups, e.g. complete forms, but please do not go below 15% of pupils.

2.2 Fill in a checklist for each pupil. Put a tick in the appropriate box for each time. If you feel that your decision is based on very little information you could indicate this by using a question mark instead of a tick. This is preferable to leaving an item blank but clearly if you are completely unable to make a decision on a particular item a blank can be left.

2.3 After completing the checklists for all the pupils in the target group (of about 20%) construct the list of the most able 10% using the checklist information to supplement any other information that you wish to use. The exact way in which this is done will inevitably vary from teacher to teacher. Broadly we would ask you to go through two stages:-

i) to consider each ofthe pupils for whom checklists have been completed in turn and to decide whether or not they should appear on the 10% list;

ii) after this initial sort to take a second look to bring the 10% group to the correct size.

> As a guideline for use in this task we would suggest that a pupil who gains more ticks in the boxes indicating high ability than is the average for the 20% group that you have worked on should be seriously considered for membership of the final 10% group. This is especially so if his/her ticks are well spread among the various categories of the checklist or are on items that are particularly relevant to the sort of work that you do.

2.4 To enable us to assess the value of the items on the checklist we need to have some indication of how they perform when the lists are used on average and low ability pupils. We have therefore included some lists on coloured paper which have the names of some pupils chosen at random from your class lists on them. We would be very grateful if you could also fill these in and return them us us. You do not, of course, need to consider these extra pupils for membership of your top 10% group.

PLEASE RETURN ALL CHECKLISTS TO US WITH YOUR FINAL LIST OF THE TOP 10%.

STAGE 1

20% List

1. Make a list of roughly the top 20% of pupils in your subject. This should be done jointly by all teachers who are involved in teaching these pupils.

2. The figure of 20% may be varied to suit your situation but please don't go below 15%.

3. You need not exercise great care in making the borderline decisions for this group.

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

STAGE 2

1. Each teacher should then fill in a checklist for any pupil on this list whom he teaches.

2. This is probably best done individually - perhaps over a period of two or three weeks.

3. Teachers should also fill in the coloured checklists on the named pupils. These need not be considered in Stage 3.

4. All teachers should then meet together for Stage 3.

STAGE 3

1. At this stage, jointly decide, on the basis of the checklists and any other information that you wish to use, which pupils from the original 20% should be on the final 10% list of pupils showing the greatest potential for future achievement in your subject.

2. Consider each pupil in turn and decide whether he/she should be in the 'top 10%' without too much regard to the numbers of pupils in this list.

3. You may then need to go through the 'top 10%' a second time to bring the group to the right size.

> As a guideline for use in this task we would suggest that a pupil who gains more ticks in the boxes indicating high ability than is the average for the 20% group that you have worked on should be seriously considered for membership of the final 10% group. This is especially so if his/her ticks are well spread among the various categories of the checklist or are on items that are particularly relevant to the sort of work that you do.

Top 10% Group

....

....

....

....

....

....

....

....

....

....

Please indicate (by putting a 'S' after the name, those pupils who are in the top half of this group. A '*' can also be used to indicate any pupil whom you regard as exceptional (anyone who, if they were to continue with the subject, could be expected to make a serious contribution to the study of or application of that subject.

If there are other pupils whose exclusion from the 10% group caused you difficulty you could list them here:

....

....

....

....

PLEASE RETURN ALL CHECKLISTS INCLUDING THE COLOURED ONES TO US.

Please complete the following two questions:

1) What information was used in addition to the checklists in producing the final IOE list?

Teachers' own records ☐

School exam results ☐

School reports ☐

Objective tests ☐

Results of 1st round of identification ☐

Other (please specify) .... ☐

....

....

2) Which teachers were involved;

....

....

....

....

....

COMMENTS

If you wish to do so, please add any comments on the checklist items (ambiguity of wording, difficulty in observing the behaviour described, items omitted etc.) and on the general way in which the lists were used.

ENGLISH CHECKLIST (YEAR 1)

NAME OF PUPIL ....

WRITING

1. Technical Control

i) Only writes simple sentences/basic syntactical structures	☐☐☐☐☐	Writes complex sentences/ shows syntactical sophistication
ii) Punctuates inaccurately	☐☐☐☐☐	Able to manipulate punctuation to convey intended meaning

2. Reader Awareness

Fails to give reader adequate background information in own writing	☐☐☐☐☐	Aware of information needs of potential reader; provides full context in own writing

3. Variety of Register

Confined to own voice	☐☐☐☐☐	Able to adopt variety of registers/styles as suitable

4. Humour/Irony

Weak sense of humour in writing (conventional or functional)	☐☐☐☐☐	Has sense of irony able to make connections revealing humour/absurdity of a situation

5. Vocabulary

Limited resources	☐☐☐☐☐	Experiments with words/sounds and new meanings

6. Originality

Restricted to predictable plots/situations in fiction	☐☐☐☐☐	Strives for fresh, inventive plots/situations in fiction

7. Empathetic Imagination

Does not describe or imagine outside own view point.	☐☐☐☐☐	Able to imagine and write from wide variety of view points

ORAL

(High oral ability refers here to a pupil's use of speech in the classroom to discover, investigate and express ideas and feelings - whether in class discussion, group talk or one-to-one conversation - and the extent to which the pupil listens and responds to what others say/have said.)

8. Oral Ability

i) (talking)	low	☐☐☐☐☐	high
ii) (listening)	low	☐☐☐☐☐	high

(In order to relate the quality and quantity of a pupil's talk, please indicate how generally talkative the pupil is in the classroom.)

9. Talkativeness

Low ☐☐☐☐☐ High

READING

10. Comprehension

i) Factual, Limited to extracting simple facts from a passage/passages	☐☐☐☐☐	Able to extract, select, synthesize facts from a passage/passages
ii) Meaning, Unable to detect implied subtle meanings, especially in poetry or fiction	☐☐☐☐☐	Sensitive to subtle or implied meanings

11. Personal Reading

i) Enthusiasm, Reads little by choice	☐☐☐☐☐	Is a keen and voracious reader
ii) Range, Limited in choice to familiar or easy reading material (material of immediate interest)	☐☐☐☐☐	Is willing to tackle a wide range of reading material (including unfamiliar or demanding)

12. Reading Out Loud

Reads out loud falteringly and monotonously	☐☐☐☐☐	Can read out loud spontaneously with appropriate sensitivity and intonation

13. Critical Judgement

Restricted to unexplained expressions of personal opinion	☐☐☐☐☐	Able to elaborate on opinions, by reference to text etc.

Please add below any comments of your own on the way your response to particular items was influenced by your knowledge of the pupil's character, (if you feel this is necessary).

FRENCH CHECKLIST (YEAR I)

NAME OF PUPIL ....

The boxes 1-4 on the checklist are to be ticked according to the following key:-

1) This item described this pupil's behaviour on all or almost all occasions.

2) This pupil frequently behaves in the way described.

3) This pupil sometimes behaves in the way described but has difficulty in maintaining this level of work.

4) This pupil rarely or never displays the behaviour described.

The one word headings to the column could be misleading on their own and should be regarded just as memory aids.

	1 Always	2 Often	3 Sometimes	4 Rarely
1. Flexibility accepts validity of concepts unfamiliar in English (adjectival agreement, altered word-order etc.).				
2. Rule Acceptance is able to embrace a system of rules as an essential basis for learning.				
3. Memory shows clear evidence of good memory.				
4. Auditory learning ability (shows clear utilisation of aurally presented information (ear-minded).				
ORAL SKILL				
5. Aural skill (i) able to discriminate between English & French.				
(ii) able to assimilate foreign sounds.				

DOCUMENT RESUME

ED 229 359 SP 022 234

AUTHOR	Denton, Cliff; Postlethwaite, Keith
TITLE	The Identification of More Able Pupils in Comprehensive Schools. Final Report. Part II.
INSTITUTION	Oxford Univ. (England). Oxford Educational Research Group.
PUB DATE	Dec 82
NOTE	294p.; For related document, see SP 022 233.
PUB TYPE	Information Analyses (070) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC12 Plus Postage.
DESCRIPTORS	Ability Grouping; *Ability Identification; *Academically Gifted; Checklists; Foreign Countries; *Informal Assessment; Inservice Teacher Education; Secondary Education; Secondary School Teachers; Skill Analysis; Standardized Tests; *Student Characteristics; Student Evaluation; *Teacher Attitudes; *Test Reliability
IDENTIFIERS	England

ABSTRACT

In the second year of a project investigating the ability of secondary school teachers to identify high-ability students, two questions were addressed: what student characteristics influenced teachers' judgments, and why checklists appeared to have little impact on teachers' judgments. A structured approach was developed to study student characteristics to which teachers were sensitive when making judgments on ability. This approach sought to determine the scope of teachers' knowledge of students, and how this knowledge corresponded with judgments based on students' test scores. Checklists were modified to make them more workable. Extensive classroom observations were made to establish if relevant clues to ability could emerge from day-to-day activities. A number of behavioral and motivational influences on teacher judgments, as well as those based on student test scores, were observed. It was concluded that teachers' assessments were sufficiently high for them, combined with refined checklists, to be considered relevant. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED229359

OXFORD EDUCATIONAL RESEARCH GROUP

THE IDENTIFICATION OF MORE ABLE PUPILS IN COMPREHENSIVE SCHOOLS

by

Cliff Denton

and

Keith Postlethwaite

FINAL REPORT

PART II

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. C. J. Denton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented to the
Department of Education and Science
December 1982.

CHAPTER NINE

The Second Year

If we were to leave the research at this point, we could claim to have achieved important insights into the success of the teacher-based identification process as we could expect it to proceed in many classrooms. We could encourage teachers to be more confident in their judgements than they might have been without our work; though we would add that there was evidence to suggest that there were some inaccuracies associated with teacher-based identification in a considerable number of cases. These inaccuracies seemed to be particularly prevalent in some subjects. There was also evidence that overemphasis on some inappropriate characteristics seemed to lead to bias in teachers' judgement. Finally, we have shown that checklists did not have any measurable effect on the identification. Merely asking a teacher to try to recognise appropriate criteria did not work - despite the fact that teachers were quite co-operative in the exercise.

Two major questions were apparent at the end of this work. One concerned the pupil characteristics that influenced teachers judgement. As we have mentioned, we had some interesting insights from the work which was reported in Chapter Four, but these were necessarily limited to information that was related to our test battery. We also had some tantalising clues from our interviews with teachers (Chapter Seven) about the kinds of characteristics that led them to under- and over-estimate potential. However the interviews were deliberately conducted in a very open ended way, and something with a more precise structure was called for if any firm conclusions were to be drawn.

The other question which was unresolved at the end of Year One was that of why the checklists, used in the way in which we had asked teachers to use them, had been of such little value in refining teachers' judgements. One possibility was that the detailed clues to pupils' ability which were referred to by the lists could not be picked up by teachers in the restrictive environment of the classroom because the pupils were not provided with the kinds of work which would enable them to display such clues. Another alternative, however, could be that the way in which the checklists were used was inappropriate.

We would suggest, then, that we cannot provide advice about how

3

improvements in teacher-based identification can be made without further study. It is of considerable importance to distinguish between alternative plausible hypotheses, and to establish whether some of the reasonable hypotheses that we have put forward can in fact be developed into a well-supported structure.

We will see, in the next few chapters, 3 ways in which we have attempted to achieve these aims.

First we developed a much more structured approach to the study of the pupil characteristics to which teachers were sensitive when making their judgements on their pupils' abilities. From this we could discover the scope of teachers' knowledge of pupils and an understanding of which aspects of their knowledge were associated with judgements that were in line with judgements based on the pupils' test scores, and which were distractors, leading teachers to over or underestimate pupils in comparison with their test scores.

The details of this method are discussed in Chapter Ten and the results are to be found in Chapter Eleven.

Secondly we gave checklists a further chance. By modifying their methods of use through negotiation with the teachers, we sought to make them more workable.

Thirdly we made extensive classroom observations to see if we could establish if the relevant clues to ability could emerge from the day to day activity of pupils in class. It may have been that because of the many limitations of time and opportunity it was not possible for teachers to have the detailed knowledge of children that is required if checklists are to be of help. This intensive activity was possible in just one subject.

The more detailed study of checklists is described in Chapter Twelve.

Thus we have attempted to come away from the computer output into the classroom so that we could add the necessary evidence which would allow the reader to make a more balanced judgement about the problems and possibilities associated with teacher-based identification strategies.

4

CHAPTER TEN

An Approach to the Understanding of Teachers' Views of Pupils based on Personal Construct Theory

In Chapter Seven we described a series of interviews with teachers through which we gained some interesting insights into the kinds of pupil characteristics that they appeared to notice. We also reported that we were able to form a view that some of these characteristics seemed to be associated with teachers' overestimation of pupils' potential, while others seemed to be associated with underestimation of pupils. In that earlier chapter we stressed the tentative nature of these insights. We drew attention to the fact that we could not be certain that we had assembled a complete picture of the range of characteristics that were noticed by teachers. We also pointed out that we had no basis on which to estimate whether the comments made by a teacher in speaking about one individual pupil were part of that teacher's way of looking at all pupils, or whether the comment was very specific to the individual pupil who was being discussed. To these, and the other limitations of the interview technique, we would add one further comment: the procedure generated interesting hypotheses concerning the factors which were associated with under- and over-estimation of potential, but it provided no grounds on which the validity of these hypotheses could be tested. This chapter, and the next, describe the background, and then the results of a more detailed exploration of these issues.

The theoretical basis that we have adopted for this more detailed exploration of the way in which teachers perceive their pupils is that of Personal Construct Theory (Kelly, 1955). Pope and Keen (1981) draw attention to the relevance of this theory for the study of many kinds of problem in education and point out that little effort has as yet been expended on the particular topic of teacher assessment of students despite the fact that, in their view, the theory had much to offer in this area.

In this context of teacher assessment, what the theory suggests is that the teacher makes sense of his interactions with a wide range of pupils by setting up a number of reference axes (or constructs) in his mind and placing each pupil at some point in the multidimensional space defined by these axes. For example: a given teacher may think in terms of a three dimensional space consisting of the axes 'quiet-noisy', 'neat-untidy' and

5

'able-not able'. He may then place an individual pupil at the left hand extreme of the axis labelled 'quiet-noisy', at the centre of the axis labelled 'neat-untidy' and near the right hand end of the 'able-not able' axis. This placement of the pupil in the teacher's three dimensional map enables the teacher to make sense of his previous experience of this pupil and to predict how the pupil may behave in new situations (e.g. since the pupil is 'coded' as quiet and not very able, he will not be expected to make many insightful contributions in a new situation such as a cut and thrust debate about a complicated issue).

The axes set up by an individual teacher may or may not be orthogonal to each other. If they are orthogonal, the placement of a pupil on one axis can be entirely independant of his placement on the other axes. If they are not orthogonal, the teacher is using a map in which one characteristic is taken as being correlated with another (If our teacher, in the example quoted above, regarded most quiet people as able, the 'quiet-noisy' and the 'able-not able' axes would not be orthogonal in that teacher's map.)

Using such a model, the problem of understanding how teachers go about the business of identifying pupils as more able can be thought of as consisting of several separate steps: first we need to know the nature of the individual axes - or constructs - that the teacher uses; secondly we need to know the structure of the relationship between these individual constructs (e.g. which constructs are orthogonal); thirdly we need to know where the teacher places each pupil in the map defined by these constructs; and finally we need to know how the teacher interprets the pupil's position in this map in trying to come to an overall decision about whether or not that particular pupil should be classified as more able. We will consider some of these stages later, but we will first continue with some further general comments about the decision to use Personal Construct Theory as the basis of this part of the work.

We suggest that the theory provides a particularly useful context for our detailed study of the way in which teachers come to a decision that a particular group of pupils have high ability in a given subject. This can be seen, first, from the fact that the theory is essentially a Personal Construct Theory: i.e. it demands that we look at the nature of the constructs used by each individual teacher, and it therefore allows us to continue to deal with the variety that we have already found to be present in the way in which different teachers look at their pupils. Secondly, by encouraging us to look at the place each pupil occupies in the teacher's construct space, it provides a framework for the collection of quantitative data which, nevertheless, is still related to the teacher's own personal system of mapping. This quantitative data is more amenable to analysis than, for example, the qualitative descriptions of pupils elicited in the interviews. Thirdly, the theory draws our attention to the fact there there may be relationships between the teacher's judgement of pupils on one construct and his judgement of those pupils on some other constructs. This is consistent with our wish to explore more fully the nature of the association between particular characteristics noticed by the teacher, and that teacher's overall assessment of the potential of a pupil.

Having established the relevance of the personal construct viewpoint to our present concerns it is useful to consider where the adoption of such a theory places us in relation to the range of views which are held about the nature of person perception. It must immediately be admitted that it places us firmly in the school of thought which regards person perception as a matter of manipulation of component pieces of information which ultimately leads the perceiver, rather in the manner of Sherlock Holmes, to some overall judgement about the perceived. In contrast, there is a second broad school that regards person perception as essentially an intuitive process. This position is described by Cook (1979) as one in which people understand others 'by an act of intuition encompassing their personality as a whole'.

This school of thought, which puts forward similar views to those advanced by Gestalt psychologists, suggests therefore that we respond to the 'meaningful whole' of others and that this 'whole' can never be adequately represented as a sum of component parts. This is the theoretical framework into which ideas such as 'feminine intuition' can be fitted, and in some circumstances it certainly seems to ring true. It accounts, for example, for the situation in which an interviewer 'just knows' that the candidate is going to be right for the job.

Despite the superficial attractiveness of the intuition-based view of person perception, it has a major shortcoming compared with the inference-based model which we have chosen to adopt. This shortcoming arises out of the fact that it is extremely difficult to see what kind of detailed investigation could (or should) be undertaken if all that we can say about

7

teachers' perceptions of pupils is that they 'just knew'. It is therefore a model which leaves one with no indication of which paths to follow in trying to help teachers to be able to identify able pupils more accurately. This is the basis of our decision to discard this intuition-based model. In doing so we can draw support from the views expressed by Cattell (1937) who described the intuition-based model as obscure and unscientific and argued that it:

> contributes nothing except a magic word the purpose of which is to cry 'Halt' to the irreverence of further investigation.

We therefore suggest that there is considerable justification for our recourse to the inference-based model. However it would be wise to draw attention immediately to the debate which exists about the precise nature of the process through which the inferences are made. In doing so we hope to provide some theoretical basis for the kinds of analysis which we have undertaken and to draw attention to one further advantage of the Personal Construct Theory approach (i.e. that it does not commit us to one particular view of the process of inference but that it provides, instead, data which can be analysed in a variety of ways, drawing upon different theoretical models as appropriate).

One much discussed model of the process of inference is that offered by Sarbin (1944) who makes an analogy with the logic of the classical syllogism:

All X are Y.
Z is X.
So Z is Y.

Sarbin suggests that this model has to be modified to the extent that the major, and sometimes also the minor premises are, in the person perception context, expressed in probabilistic terms. For example:

80% of sports car drivers are aggressive.
John is a sports car driver.
So there is a 0.80 chance that John is aggressive.

Considerable debate has been generated by this model (Meehl, 1954; Sarbin et al, 1960; Meehl, 1961). In summarising this debate, Cook (1979) makes the point that 'nothing follows logically from the major premise "Some X are Y"'. This view would seem to be consistent with that of Allport, which is quoted - with, it seems, some sympathy -by Meehl (1954). Their argument rests on the idea that it is a nonsequitur to conclude that 'there is a 0.80 chance that John is aggressive', for, they say, we know that the real position is that John is either aggressive (in which case there is a chance of 1.00 that he is aggressive), or he is not (in which case the chance is 0). These are the only possible positions and in both of them the figure of 0.80 has no real meaning. They suggest that the information provided is not sufficient to enable us to pass any judgement on John's aggressiveness on the basis of reasoning of this kind. The implication in this line of argument is that, since we do come to conclusions about people on the basis of such data, we must actually structure our reasoning in some other way. We will present alternative models (including that due to Meehl) shortly, but we feel that it is worth looking a little more closely at the substance of this argument before moving on.

While we accept that, as far as it goes, the point of view expressed in the previous paragraph is sound, we would suggest that it leads one to undervalue Sarbin's model. Surely the point is that, in the absence of any of the further information necessary for the certain classification of John, the most sensible thing that we can say about him is that he (like 80% of sports car drivers) will be aggressive, and that the figure of 80% gives us a basis for assessing the odds that we will be right in making such a judgement. We would therefore suggest that the reasoning may run as follows:

80% of sports car drivers are aggressive.
John is a sports car driver.
So in the absence of further data I will assume that John is aggressive
And the odds are 4:1 that I shall be proved right.

This model has moved some way from the classical syllogism but it is still a model based on the categorisation of the individual, and on inferences drawn from what is known about the category into which the individual is put.

We turn now to Meehl's own alternative model of the process of inference which is essentially a correlation model. Meehl (1961) proposes that the perceiver does not think of placing the perceived person into categories

9

and of indulging in something similar to syllogistic reasoning on the basis of that categorisation. Rather, he suggests, the perceiver reasons on the basis of the degree to which the perceived person has some property 'x', and on the basis of the strength of the correlation between this property 'x' and some other property 'y', so as to come to some view on the degree to which the person can be expected to have this second property 'y'.

Because, as we have suggested above, there is no watertight logical reason to discard Sarbin's model, we suggest that the choice between the Sarbin and Meehl models is by no means straightforward when one is considering the simple situation in which a person is reasoning on the basis of a single cue. However Meehl makes a strong case that his model, when generalised from simple correlation to multiple regression, is far more successful than the syllogistic model in describing the situation in which a perceiver is trying to handle a large number of partially correlated cues. Indeed it is clear that any attempt to express such a problem as a series of syllogisms would be enormously complex.

It would seem that the data which we have collected within the personal construct framework (consisting as it does of ratings of pupils on each of their teacher's constructs) is very amenable to analysis in terms of this generalised version of Meehl's correlational model for the concept of 'the degree to which the pupil has some property 'x'' is clearly represented in such data. However, as we hinted earlier, one advantage of the personal construct approach is that it does not limit us to this model of the process of inference. Cook (1979) describes a number of alternative models for the way in which people handle multiple cues. He suggests that people may add, or average information on different aspects of another person to come to an overall judgement about that person, and that inconsistent information may be discounted. He suggests that in the process of discounting some parts of the set of information, positive information tends to be overlooked more readily than negative information.

He also suggests that people may make use of favourite cues and of disqualifying cues. In our context, this last suggestion would imply that the teacher may be sensitive to a particular pupil characteristic that he/she regards as so firmly linked to high ability that a pupil displaying such a characteristic will almost always be described as more able whatever any further information on the pupil might indicate (such a characteristic would be operating as a favourite cue for that teacher).

10

Similarly the teacher may be sensitive to another characteristic which is so firmly linked in the teacher's mind with low ability that pupils displaying this characteristic will almost always be denied membership of the more able group whatever other information is available (such a characteristic would be operating as a disqualifying cue).

The concepts of favourite and disqualifying cues (which we would see, incidentally as more consistent with the syllogistic, rather than the correlational model of the process of inference) are of particular interest to us as we already have some indication from the general interviews with teachers that they seemed to use some characteristics in this way. If this could be confirmed by more detailed study valuable insights could be gained into why teachers overlook some pupils whom the tests suggest are of high ability, and why they choose others whom the tests suggest are not of high ability. Once the nature of these problems is understood, steps could clearly be taken to help teachers to be more accurate in their overall judgements.

We have therefore given some thought to the ways in which personal construct data could be used to investigate these ideas. We have considered three major topics: first how the concepts of favourite and disqualifying cues could be assimilated into the personal construst framework; secondly how one might try to elicit the nature of the cues which a teacher might be using as favourite, or disqualifying cues; and thirdly how one might test the hypothesis that these cues were actually being used in these ways. In the rest of this chapter we discuss the first two points in some detail. The third will be discussed in the next chapter, where the results of this work are presented.

The relationship between the concepts of favourite and disqualifying cues, and constructs is essentially a simple one, but we will discuss it here as it allows us to introduce some terminology that will be helpful elsewhere in the argument. We begin by discussing the disqualifying cue.

We suggest that the idea of a disqualifying cue may be helpful in trying to understand why teachers overlook some pupils who, according to the tests, have high potential for the subject. We could call such pupils 'misses' and could label them as ($T\bar{P}$) (i.e. nominated by the test (T) but not by the teacher - i.e. not of high perceived potential ($\bar{P}$)). Because we wished to continue to take account of test error in identifying pupils who fitted this specification we decided that we would not count any pupil who fell within the top 10% group as a (T) pupil (a pupil nominated by the test), but would restrict this label to pupils who were members of the A Group (see Chapter Three). The idea that a teacher is influenced by a disqualifying cue in forming an overall judgement about such pupils suggests that he/she may have a single construct (or group of constructs) which have one pole which is so firmly a negative correlate of that teacher's concept of high potential that pupils placed at this pole will tend to be overlooked (labelled as $(\bar{P})$) irrespective of their position in the rest of the teacher's construct space.

Similarly, the idea of a favourite cue may be helpful in trying to understand why teachers choose some pupils who, according to the test, do not have high potential for the subject. We could call such pupils 'false hits' and could label them as $(\bar{T}P)$ (i.e. not nominated by the test $(\bar{T})$ but nominated by the teacher - i.e. of high perceived potential (P)). Again, to take account of test error we decided to restrict the label $(\bar{T})$ to those pupils who fell into the C Group. The idea that a teacher may be influenced by a favourite cue in forming judgements about these pupils suggests that he/she may make use of a single construct (or group of constructs) which have one pole that is so closely associated with that teacher's view of high potential in the subject that pupils placed at this pole will tend to be chosen (labelled as (P)) irrespective of their position in the rest of the teacher's construct space.

There is no guarentee that disqualifying and favourite cues will lie at opposite poles of the same construct (e.g. a teacher's disqualifying cue may be 'uncouth behaviour' but his/her favourite cue may be 'good at arithmetic') so care must be taken to ensure that both kinds of construct can emerge from a personal construct interview.

Before turning to our second problem of how the constructs that involved favourite and disqualifying cues could be elicited, we will give a very brief description of the usual methodology employed in personal construct studies.

The classical approach to the problem of how to find out what kinds of constructs an individual uses in structuring his or her experience of a particular group of elements (e.g. people in the individual's family, people he/she works with, types of music, styles of writing or, as in our case, pupils taught by the individual) is the approach known as triadic elicitation. In this procedure a list of relevant elements is produced and groups of three elements from this list are constructed. Each group of three elements (each triad) is then presented to the individual who is asked to say in what way two of the elements are alike, and in what way the third is different from the other two. For example, in a study of reactions to types of music we might present a triad consisting of Offenbach, Gilbert and Sullivan and Wagner operas. From such a triad we might elicit the comment that Offenbach and Gilbert and Sullivan operas are light, whereas Wagner's operas differ from the others by being serious. We would then know that one construct (or reference axis) used, in relation to music, by the person to whom we had offered the triad, was the construct: 'serious -light'. We would continue by offering other similar triads, so that we could begin to build up a full picture of how our subject thinks about types of music. This procedure, together with other aspects of the use of Personal Construct methodology, is described briefly by Hall (1978), and is discussed in detail by Fransella and Bannister (1977).

In preparation for the discussion of how this procedure must be modified to suit our main purpose it is useful to note that personal construct studies are often used to provide descriptions of the way in which people view some aspect of their experience, or as an aid to diagnosis, for example in the treatment of psychiatric patients. In such circumstances, the overriding consideration governing the elicitation procedure is that it should provide a representative map (or ideally a complete map) of the whole range of constructs used by the individual in relation to a particular kind of element. This consideration implies that the triads we present to the individual should be representative of all possible triads that could be constructed from the relevant list of elements; and this, in turn implies that the most suitable way of choosing the three elements for each triad is on the basis of random selection. This advice is clearly stated by Yorke (1978) who advises that:

> without any strong justification for choosing to use particular triads it is probably preferable to assign elements randomly.

Certainly one purpose of our personal construct work was to provide a description of the kinds of pupil characteristics that were noticed by teachers so that we could extend the general insights gained from the interview study; but this was largely incidental to the second, and more

13

important purpose, which was, as we have explained above, to explore the idea that some constructs might be associated with favourite cues, and others with disqualifying cues.

In connection with the first purpose a representative mapping was required and, therefore, the presentation of randomly chosen triads was called for. However the same considerations do not apply in connection with the second purpose. The over-riding consideration there, was that of encouraging the teacher to mention the constructs that were associated with any favourite and disqualifying cues that he/she might use. Clearly the appropriate constructs would only emerge if a $(T\bar{P})$ pupil (one who was overlooked despite high test score) or a $(\bar{T}P)$ pupil (one who was chosen despite low test score) was presented, in a triad, with appropriate contrasting pupils.

By reference to Chapter Three the reader will be able to see that because of the considerable degree of match between test and teacher judgement of pupils, there were relatively few pupils who could be labelled as $(T\bar{P})$ or $(\bar{T}P)$ in any one subject in any one school. Therefore, only if large numbers of triads were presented to each teacher, could one rely upon the random selection of pupils for each triad to generate the appropriate groupings. This recourse to large numbers of triads was itself rather unattractive, partly because it would have been unacceptably demanding on the time of the teachers, but more seriously because the presentation of large numbers of triads would have been exhausting for both parties involved in the elicitation interview (Pope and Keen, 1981). As a result, the long interviews might be expected to become more and more unrewarding as they continued. We were therefore drawn back to a position not unlike that adopted by Kelly himself, in his seminal work on Personal Construct Theory (Kelly, 1955, p.273 et seq). He suggested using, and gave examples of, carefully chosen triads which could be offered to elicit constructs of a particular kind. Of course, as the kinds of constructs in which we were interested differed from those of which Kelly was writing, the details of the formulae for the construction of these special, or 'critical' triads were necessarily different in our case. Nevertheless the principle is the same. This idea has also been suggested by Easterby-Smith (1981) though he provides nothing more than the hint that it might be useful to present triads which will 'bring out the greatest contrast in the elements available'.

As an example of the kind of critical triad that was relevant to our work, consider the following group of pupils $(T\bar{P}; T\bar{P}; TP)$. These three pupils had similar high scores on the test measure of potential and so were all classified as (T). One also had high perceived potential (P), but two were not regarded as members of the top 10% group by their teacher $(\bar{P})$. If, in the elicitation interview, the teacher grouped the pupils thus $(T\bar{P}; T\bar{P})$ and (TP), one might expect that the resulting construct would be an important part of the way in which that teacher was led towards making errors of under-estimation, and that it contributed to the 'misses' that he/she made. One might also suspect that one pole of this construct (that describing the pair of $(T\bar{P})$ pupils) might have been operating as a disqualifying cue for that teacher. Clearly other distinctions between the pupils in the triad might override this distinction based on test and teacher perceptions of potential and one would not, therefore, expect all triads of this type to be split by the teacher in the way which was suggested above. Nevertheless when such a split was made, it was possible to make the kind of classification of the construct that we have indicated above.

A very important point emerges from this paragraph. Not only does the presentation of carefully designed 'critical' triads encourage the teacher to mention the particular kinds of construct in which we were especially interested, it also allows us to make an *a priori* classification of the resulting construct. The appropriate constructs might emerge if enough randomly designed triads were presented but no *a priori* classification would then be possible. Because the 'critical' triad approach gives rise to this possibility of forming classifications of triads that can later be tested to see if it holds up, it effectively shifts the technique from one of description and diagnosis to one of *experimental* exploration of the topic. We believe that this represents a fairly novel way of putting personal construct methodology to work.

Having established the idea of 'critical triads' we turn now to the question of how many different kinds of critical triad can be designed with the kind of data which we have in our study. For this part of the work we can regard our problem as one involving two-way classification (pupils are labelled as more able or not by the teacher, and as more able or not by the test). There are therefore four different types of pupil (TP), $(\bar{T}P)$, $(T\bar{P})$ and $(\bar{T}\bar{P})$. We can clearly make up four types of triad in which all three pupils are the same, and four types in which all three pupils are different from each other (e.g. TP, $\bar{T}P$,$T\bar{P}$). In each of these cases a priori classification of the constructs which emerged from the triads would be impossible. However constructs emerging from the other twelve triads (in which one pupil differed from the other two) could be classified on an a priori basis. These triads are listed in Table 10.1, together with the appropriate classification.

TABLE 10.1

POSSIBLE CRITICAL TRIADS

Triad type A1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, TP

Triad type A2: TP, TP, $\bar{T}\bar{P}$

These triads give rise to constructs on which the teacher's judgement categorised the pupils in the same way as the test.

Triad type B1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, $\bar{T}P$

Triad type B2: $\bar{T}P$, $\bar{T}P$, $\bar{T}\bar{P}$

These triads give rise to constructs on which the teacher separated pupils with similar low test score. The pole at which $\bar{T}P$ was placed may have been a favourite cue.

Triad type C1: TP, TP, $T\bar{P}$

Triad type C2: $T\bar{P}$, $T\bar{P}$, TP

These triads give rise to constructs on which the teacher separated pupils of similar high test score. The pole at which TP was placed may have been a disqualifying cue.

Triad type D1: $T\bar{P}$, $T\bar{P}$, $\bar{T}P$

Triad type D2: $\bar{T}P$, $\bar{T}P$, $T\bar{P}$

These triads give rise to constructs on which the teachers' perceptions lead to reverse assessments from those based on the test. Favourite and disqualifying cues may be found at the opposite poles of these constructs.

Triad type E1: $\bar{T}\bar{P}$, $\bar{T}\bar{P}$, $T\bar{P}$

Triad type E2: $T\bar{P}$, $T\bar{P}$, $\bar{T}\bar{P}$

In these triads $\bar{T}$ and T are distinguished by the teacher but the characteristics displayed by the $T\bar{P}$ pupils are undervalued as indicators of high potential. The $T\bar{P}$ pole may again be a disqualifying cue.

Triad type F1: TP, TP, $\bar{T}P$

Triad type F2: $\bar{T}P$, $\bar{T}P$, TP

In these triads T and $\bar{T}$ are distinguished by the teacher but the characteristic displayed by the $\bar{T}P$ pupils is undervalued by the teacher as an indicator of low potential. The TP pole may again be a favourite cue.

Throughout the table the construct classification only applies if the triad is split so that the third member is identified as the 'odd one out'.

It must be recognised that an elicitation procedure based only on these critical triads may produce a distorted grid, limited by our own perception of the nature of the problem (Yorke, 1978). Our major response to this criticism is that we have traded a representative survey of teachers' constructs for an experimental study of constructs which behave as favourite and as disqualifying cues. We also note that, through critical triads A1 and A2 we get insights into the nature of the constructs which lead teachers to make classifications of pupils which are in line with those made by the tests. However, to go some way to overcoming this criticism in a more direct manner, some triads made up of randomly chosen pupils were included in our elicitation interview together with triads fitting the specifications listed in Table 10.1. In presenting the triads, critical and randomly chosen triads were intermixed, and the overall order of presentation was chosen at random.

Names of pupils to form the critical triads were selected in the following way. The pupils taught by a given teacher were each labelled as TP, $\overline{T}$P, etc. When a pupil of a particular kind (e.g. TP) was needed in a triad, a name was chosen at random from the set of pupils of that kind who were taught by that teacher. When triads suiting all of the specifications in Table 10.1 had been constructed, they were listed in the order of presentation. They were then inspected and some changes were made, if necessary, to avoid repetition of pairs of names in successive triads. (Bender (1974) found that there was a significantly greater tendency to produce important constructs when successive triads were varied by two elements at a time, rather than one at a time.)

The intention was that the elicitation interview should present each teacher with fifteen different triads. Some teachers, who felt that they had exhausted their repertoire of constructs before this point had been reached, actually handled fewer than fifteen triads. Others dealt with several more. The time available for the interviews (normally something of the order of 90 minutes) enabled this scale of activity to be conducted in a busy, but not unduly rushed fashion. It is interesting to note that all of the teachers who went beyond fifteen triads felt that they had nothing extra to offer after dealing with at most five additional triads. It seems that it would indeed have been of little value to attempt to work with very large numbers of randomly chosen triads.

In the next chapter we present the results obtained from this study.

The work was done in detail in physics and on a smaller scale in the other subjects. We therefore present the physics results first to give a clear picture of the various stages of the analysis and interpretation of the findings. We then summarise the outcome of the work in the other subject areas.

CHAPTER ELEVEN

Results of the Personal Construct Study

In this chapter we give an account of the findings which have arisen to date, out of the personal construct work that was described in the previous chapter. It must be emphasised that this has proved to be a very fertile line of enquiry and that the resulting data lends itself to a much more extensive analysis than we have been able to undertake within the constraints of the present project.

The main aim of the work which has so far been completed has been to gain insights into the kinds of constructs which teachers use in assessing pupils, and to indicate:

i) which of these constructs seemed to be associated with accurate overall judgements of pupils ability;

ii) which seemed to have poles that operated as favourite cues, and were therefore associated with over-estimation of ability;

iii) which seemed to have poles which operated as disqualifying cues, and were therefore associated with an under-estimation of ability.

Some data relating to these main aims is presented for all four subjects. However, the demands on time made by other parts of the project led us to undertake the majority of the fieldwork in physics. We have therefore taken the analysis of the physics data rather further than the analysis of the data in the other subjects. We will therefore begin by reporting this work in physics. It is perhaps worth emphasising that even in this subject, there is much more that could be done (and that we hope will be done) in the future.

11.1 Basic results in Physics

In physics, 182 triads were presented in all to 11 teachers in 7 schools: 140 of these triads were critical triads; 42 were randomly chosen ones. In the case of 19 triads (12 critical triads, and 7 randomly chosen ones) the teacher found it impossible to divide the pupils into a group of two, and a contrasting odd-man-out. It is interesting to note that 8% of the

20

critical triads proved to be impossible for the teacher to split, whereas 17% of the random triads suffered this fate. This is not surprising, in that the critical triads were specifically chosen to contain pupils who differed in some way, but it does suggest that the use of critical triads did, as predicted, add to the efficiency of the elicitation procedure[1]

Of the 128 critical triads that were split by the teachers, 80 (62.5%) were split as one might have expected from the design of the triad. It is difficult to draw any firm conclusions from this figure, as not all of the triad divisions made by teachers can be thought of as independant of one another. (For example, the same pupils sometimes appeared in several of the triads which were offered to a given teacher.) However, the number of triads that were divided in the expected way, was nearly twice the number that one would expect to be divided in that way if the teachers' divisions of the triads had been completely uninfluenced by the design of those triads. In most cases, this high number of 'expected' splits was confirmed when each teacher's choices were considered separately. The results are summarised in Table 11.1.

TABLE 11.1

CRITICAL TRIAD SPLITS ANALYSED BY TEACHER[2]

TEACHER	NUMBER SPLIT	NUMBER SPLIT AS EXPECTED
03	10	9
04M	10	7
04H	12	10
05	11	6
07G	15	8
07B	12	7
08	9	6
09	14	9
10	9	5
11	13	4
12	13	9
Overall	128	80

For comparison, the split of random triads was also analysed. There was, as one would expect, no evidence that the way in which these triads had been split had been biased towards the choice of any particular pupil in the triad (e.g. the first pupil).

Finally, each type of triad identified in Table 10.1 was analysed separately. The results are shown in Table 11.2.

TABLE 11.2

CRITICAL TRIAD SPLITS ANALYSED BY TRIAD TYPE

TYPE	NUMBER SPLIT	NUMBER SPLIT AS EXPECTED	SIGNIFICANCE (p)
A1	11	9	0.001
A2	10	6	0.12
B1	7	5	0.04
B2	12	8	0.02
C1	11	9	0.001
C2	12	6	0.18
D1	10	7	0.02
D2	13	7	0.10
E1	12	8	0.02
E2	7	4	0.17
F1	18	8	0.20[3]
F2	4	3	0.11

In general each teacher handled one triad of each type, so the splits made in each case within a particular type of triad, could be regarded as independent of the other cases within that type of triad. One can therefore calculate the probability of achieving exactly the observed number of 'expected splits' within each triad type, on the basis of chance alone. A simple summing of such probabilities then enables one to calculate the probability of making this number, or more expected splits on the basis of chance alone[4]. This figure can be interpreted as a significance level, indicating the significance of the deviation of the observed result from that which would have been obtained purely on the basis of random selection from the triads.

In six cases the results in this table were significant at the 0.05 (5%) level. In three further cases the probability that the observed number of 'expected splits' would arise by chance was approximately 0.1, and in no case did it rise above 0.2. Of course, when small numbers are associated with each significance test, the results (both those that indicate significance and those that do not) should be treated with caution. A change of just 1 in the number of splits made as expected, can make a great difference to the significance level[5]. However, the results do provide some evidence that teachers did seem to be influenced by the design of the triads. More often than could be expected if they were splitting the triads at random, they chose, as the odd-man-out, the pupil whose specification (in terms of T, $\bar{T}$, P and $\bar{P}$) would lead us to expect him to be chosen.

Table 10.1 provided an a priori classification of the constructs which could be elicited from particular triads if they were split in the 'expected' way. In the following tables we list such constructs under the relevant a priori classifications. We also include, in the Table 11.9, a list of constructs arising from random triads, and from triads that were not split in the 'expected' way. A short discussion of the results follows each table.

TABLE 11.3

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT	
	TP pole	$\overline{TP}$ pole
03	Good overall physics ability	Poor overall physics ability
04H	Asks perceptive questions	Does not ask perceptive questions
04M	Prepared to answer questions in class	Tends not to answer questions in class
05 (a)	Mature formal written work	Casual written expression
05 (b)	Doesn't exert self	Makes most of opportunities
07B	Able in mathematics	Not able in mathematics
07G	High ability	Low ability
08 (a)	Shows flair for subject	No flair shown
08 (b)	High performance	Low performance
09	Puts in a great deal of effort	Jogs along with adequate effort
10	High mathematics ability	Low mathematics ability
11	-	-
12 (a)	At home with subject	Finds physics difficult
12 (b)	Has natural ability	Has to try

These constructs indicate some of the dimensions on which the teachers assessed their pupils. The right hand poles of these constructs indicated characteristics which teachers noticed in pupils who were not judged, by them, to be more able. These were pupils who were also classified by the tests as not more able. In contrast, teachers offered the left hand poles of these constructs as descriptions of pupils who were assessed by the tests, and by the teacher, as more able. The left and right hand poles of these constructs could therefore be regarded as pupil characteristics to which the teachers were sensitive, and which they themselves associated with, or interpreted as signs of, high and low potential respectively. The pupils' test results confirmed the teachers' overall judgement of these pupils, and provided support for the view that the teachers' interpretation of the implications of the various characteristics as indicators of potential, was appropriate.

Present performance in the subject or in a specific part of it (i.e. mathematics), or high general ability (e.g. as exemplified by mature written work) seem to have been the main characteristics that were appropriately interpreted as signs of high potential. 'Flair' and 'natural ability' also seem to have been recognised. Construct (b) from Teacher 05 calls for some comment as at first sight it seems a little odd. It is interesting to speculate that the pupil 'Doesn't exert self' because he 'Doesn't need to exert self', in which case this seemingly reversed construct would fall in line with the two constructs from Teacher 12, and perhaps, construct (a) from Teacher 08.

In Table 11.4 we turn our attention to the constructs elicited when teachers split triads of Type B1 and B2 in the expected way.

TABLE 11.4

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	**CONSTRUCT** $\bar{T}P$ pole		$\overline{TP}$ pole
03 (a)	Absorbs day to day work	-	Doesn't absorb day to day work
03 (b)	Enthusiastic	-	Not enthusiastic
04H	Confident in practical work	-	Hesitant in practical work
04M(a)	Concentrates	-	Unable to concentrate
04M(b)	High overall ability in physics	-	Low overall ability in physics
05	Good, conforming behaviour	-	Non-conformist behaviour
07B	Good test results	-	Low test results
07G	-	-	-
08	High performance	-	Low performance
09	Applies him/herself in lessons	-	No drive in lessons
10	Conscientious	-	Very lazy
11	Good mathematically	-	Not good mathematically
12	Doesn't ask questions	-	Asks questions

On these constructs, teachers were separating pupils of similar, low test score. The pupils who were described in the terms set out under the $\bar{T}P$ pole, were pupils who were regarded as more able by their teachers despite the fact that their test scores were low[6]. An interesting feature of the descriptions under this pole - descriptions which we argued earlier might be acting like 'favourite cues' for a teacher - is that there is a clear reference to 'model pupil behaviour' (03(a), 03(b), 04H, 04M(a), 05, 09 10). Behaviour of this kind may of course be associated with high potential in some cases, but the indication from this work is that teachers may have interpreted it in this way even when the test score would suggest that this interpretation was unduly optimistic.

It is interesting to speculate on what was happening where teachers labelled a pupil with low test score as a pupil of 'high ability' (04M(b), 08). It may be that the process of interpretation of cues was happening so automatically at a non-verbal level, that the teacher was not aware that what he[illegible] labelled as 'high ability' was actually observed as, for example 'application' or 'enthusiasm'. An alternative explanation is, of course,

25

that pupils so labelled by test and teacher were pupils whose test scores were in error. Although attempts were made, by choosing pupils with extreme test scores, to reduce the impact of this test error on the triadic elicitation procedure, one must recognise that some individual pupils would have had a true potential which was considerably different from that indicated by the test score.

Table 11.5 indicates the constructs elicited from triads of Types C1 and C2.

TABLE 11.5

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT $T\bar{P}$ pole	TP pole
03 (a)	Scatterbrain in work	Well organised in work
03 (b)	Low accuracy in written work in physics	High accuracy in written in physics
04H(a)	Low quality content in written work	High quality content in written work
04H(b)	Slow to grasp ideas	Quick to grasp ideas
04H	Low overall ability in physics	High overall ability in physics
05	Not assertive	Makes presence felt (in class)
07B	Less able in written work	Able in written work
07G	Low performance	High performance
08	-	-
09	Questions have to be directed to him to involve him	Contributes without specifically being asked
10 (a)	Quiet worker	Interested in and enthusiastic for the subject
10 (b)	Reluctant to hand in homework	Homework usually in on time
11	Shows poor understanding of physics	Shows good understanding of physics
12 (a)	Finds concepts (in physics) difficult	Good grasp of concepts
12 (b)	Careless work	Careful work

On these constructs teachers separated pupils of similar, high test score. The pupils described under the $T\bar{P}$ pole were pupils who were not regarded as more able by their teachers, despite their high test scores[6].

Some interesting possible disqualifying cues can be seen at the left hand pole of the constructs listed in Table 11.5. The picture that emerges is that teachers interpreted qualities such as disorganisation (03a, 10b), carelessness (03b, 12b,), withdrawn behaviour (05, 09, 10a) and slowness

(04H, b) as indicative of low potential. Where constructs such as 'low-high performance' were used, the discussion set out in the previous section would seem to be relevant again.

Table 11.6 indicates constructs elicited from triads of types D1 and D2.

TABLE 11.6

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT $\bar{T}P$ pole	 $T\bar{P}$ pole
03	High self motivation	Low self motivation
04H(a)	Careful presentation of written work	Poor presentation written work
04H(b)	Quick to grasp ideas	Slow to grasp ideas
04M(a)	Low overall ability	High overall ability
04M(b)	Tends not to answer questions	Prepared to answer questions
05	-	-
07B	-	-
07G	Low ability	High ability
08	Committed to the subject	Not committed to the subject
09 (a)	Made good progress	Fails to show early promise
09 (b)	High general standards e.g. of politeness	Low general standards
10	-	-
11	Work improving	Started well but not maintained progress
12 (a)	Serious about work	Casual attitude to work
12 (b)	Passive in lessons	Active in lessons

The contrasts required of the teacher in handling these triads were different from those required in relation to Tables 11.4 and 11.5, so different constructs could be expected. However, the descriptions at the left hand poles can be thought of as contributing further to the list of perceived characteristics which were associated with over-estimation of potential, and those at the right hand pole to the list of perceived characteristics associated with under-estimation of potential. It is, therefore, interesting to see echoes in this table of the findings set out

28

in Tables 11.4 and 11.5.

In two of the constructs in Table 11.6 (04Ma and 07G), the teachers' judgements when dividing the triad seem to be inconsistent with the teachers' judgement when assessing overall potential (P). One can only assume that the teachers' views of the relevant pupils changed between making up the lists on which the labels P or $\bar{P}$ were based, and taking part in the elicitation interview.

It is interesting that two teachers (09, 11) described pupils with high test score but low perceived potential as 'failing to show early promise'. Since the test assessment of their potential, and the teachers implied early assessment of high promise, were in agreement, it is interesting to ask why such pupils fail to fulfil this promise and cease to be labelled as more able by their teachers. It may be that these able pupils were quick to respond to new challenges (such as a new course of study at school or a novel objective test) but also quick to disengage from these activities when they proved to be insufficiently demanding. Some of the other behaviours described under this pole (casual work, low general standards, low motivation etc.) might also arise as a consequence of such a disengagement.

In Table 11.7 we set out the constructs elicited from triads of type E1 and E2.

TABLE 11.7

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT $T\bar{P}$ pole		$\bar{T}\bar{P}$ pole
03	Lacks commitment to academic work	-	Committed to academic work
04H	Poor attitude to school	-	Good attitude
04M	Arrogant attitude	-	Accepting of advice and comment
05	Over confident	-	Doesn't assert self
07B	-		-
07G	High ability	-	Low ability
08	-		-
09 (a)	Jogs along - adequate effort	-	Doesn't put in as much effort as (s)he should
09 (b)	Not involved in lessons	-	Positive approach
10	Withdrawn	-	Outgoing
11	-		-
12	Quiet	-	Vocal

In the construct from Teacher 07G we see an example of a distinction which was clearly in line with the test based measure of potential. The pupils labelled as high ability but $\bar{P}$ may have been a pupil about whom the teacher changed his mind, or may indeed have been a borderline case for nomination as a more able pupil when that judgement was being made.

Construct 09 (a) is perhaps an interesting example of a case where the teacher misinterpreted the relationship linking the distinction which he made between the pupils in the triad which generated this construct, and high potential. One can speculate that this teacher was looking for extensive effort as an indicator of high ability and undervalued the pupils who merely 'jogged along'. This would indeed be confirmed by a separate but similar construct elicited from this same teacher and reported in Table 11.3.

It was also suggested above, that the left hand poles of these constructs might be used rather like disqualifying cues. It is interesting, therefore, to see the similarities between many of the characteristics in this table and the relevant ones in Tables 11.5 and 11.6.

30

In Table 11.8 we give the constructs elicited from Type F1 and F2 Triads.

TABLE 11.8

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT	
	$\overline{\text{TP}}$ pole	TP pole
03	Poor understanding of concepts	Good understanding of concepts
04H	Quick to grasp ideas	Slow to grasp ideas
04M	-	-
05	Mature in speech	Doesn't use verbal communication effectively
07B	Keen	Not keen
07G	High level of performance	Low performance
08	Neat presentation	Untidy presentation
09 (a)	Neat written work	Untidy written work
09 (b)	Not involved in lessons - few questions, ideas, suggestions	Positive approach - plenty of questions, ideas, suggestions
10	-	-
11	-	-
12	-	-

Constructs 03 and 09b are perhaps examples of constructs where the teacher places too little emphasis on the left hand poles as indicators of low potential. The left hand poles of most of the remaining constructs could reasonably be interpreted as 'favourite cues'. In Constructs 04M and 07G the nature of the descriptions which appear under the right hand pole seem to be difficult to reconcile with the fact that the pupils described were regarded as more able both by the test and the teacher. This is discussed more fully in Note 7 at the end of this chapter.

Finally in Table 11.9 we give the constructs from random triads and from critical triads that were split in an unexpected way.

TABLE 11.9

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT *A PRIORI* CLASSIFICATION

TEACHER	CONSTRUCT	
03 (a)	Lacking in confidence	- Plenty of confidence
03 (b)	Spontaneous interest in physics	- Lacks spontaneous interest
03 (c)	Poor practical skills	- Good practical skills
04H(a)	Asks a lot of questions	- Asks few questions
04H(b)	Very interested	- Shows little interest
04M(a)	Tendency to be silly	- Sensible attitude
04M(b)	Does not work hard	- Hard working
04M(c)	Prepared to think about physics problems	- Wants to be spoon fed
05 (a)	Independant worker	- Needs to work as part of a group
05 (b)	Uses ability	- Tends to sit back
05 (c)	Reluctant to put forward ideas	- Contributed to lessons
05 (d)	Shows ability in classwork	- Quiet in classwork
05 (e)	Willing to take part	- Not actively involved in work
05 (f)	Able in subject	- Not able in subject
05 (g)	Works unnecessarily hard	- Knows when to stop
07B(a)	Difficult to motivate	- Enthusiastic
07B(b)	Extrovert	- Not outgoing
07B(c)	Wants to succeed	- No drive to succeed
07B(d)	Poor presentation	- Good presentation
07G(a)	Underperforms relative to ability	- Performs to level of ability
07G(b)	Working class aspirations	- Middle class aspirations
07G(c)	Quiet retiring personalisty	- Forward personality
07G(d)	Minimum parental support	- Extensive parental support
07G(e)	Clerically efficient	- Not clerically efficient
07G(f)	Highly motivated	- Lower motivation

Table 11.9 is continued.

32

Table 11.9 continued.

TEACHER	CONSTRUCT	
07G(g)	Needs high degree of personal supervision	- Needs only minimum supervision
07G(h)	Exceptional interest in physics	- Not interested
08 (a)	Not precise in explanation	- Precise in explanation
08 (b)	Underperforms	- Performs as well as able
08 (c)	Lazy	- Conscientious
08 (d)	Low performance	- High performance
08 (e)	Overestimates own potential	- Realistic estimate of own potential
09 (a)	Takes an active part in lessons	- Diffident in lessons
09 (b)	Restless	- Normal conforming behaviour
09 (c)	Very competent	- Poor performance
09 (d)	Conforms to school policy	- Does not conform to school policy
09 (e)	Devoted to subject	- Not devoted to subject
10 (a)	High overall intelligence	- Low overall intelligence
10 (b)	Displays leadership qualities	- Doesn't display leadership
10 (c)	Will actively seek help	- Does not actively seek help
10 (d)	Asks questions freely	- Rarely asks questions
10 (e)	Test performance above classroom performance	- Balanced performance
10 (f)	Doesn't understand physics concepts	- Does understand physics concepts
10 (g)	Positive attitude to subject	- Negative attitude
10 (h)	Confident in his/her ability	- Not confident in ability

Table 11.9 is continued.

33

Table 11.9 continued.

TEACHER	CONSTRUCT	
11 (a)	Introvert	Outgoing
11 (b)	Lazy	Work oriented
11 (c)	Antagonistic to physics	Keen towards physics
11 (d)	Outspoken	Quiet
11 (e)	Competitive against others	Compete with own standards
11 (f)	Doesn't do homework	Does homework
11 (g)	Stands back from practicals	Gets on with practicals
11 (h)	Criticism has little effect	Criticism has effect
11 (i)	Not imaginative	Imaginative
11 (j)	Does best without being pushed	Needs to be pushed to do best
11 (k)	Very articulate	Not articulate
12 (a)	Reluctant to get on with work	Gets on with work
12 (b)	Likely to answer questions	Unlikely to answer questions
12 (c)	Haphazard in work	Methodical in work
12 (d)	Confident at subject	Lacks confidence at subject
12 (e)	Interested in physics out of school	Not interested out of school

This table completes the list of constructs elicited in our study of physics teachers, and allows us to make a general comparison between our overall results and those obtained by Wood and Napthali (1975) in their study of the constructs that teachers of geography and physics used in the assessment of the achievements of secondary school pupils. All but two of the eighteen constructs which they listed as having been mentioned by three or more of their teachers, are to be found in Tables 11.3 to 11.9 above. Most of their other constructs are either matched exactly in our findings, or find close analogies there. There would seem to be little indication, at least in comparison with that source, that major areas of our teachers' construct systems, which had relevance to their assessment of potential, were overlooked. We return to this point later.

It is also interesting to note that Wood and Napthali found that 'cognitive constructs play a central, although not overwhelming part, in the assessment of achievement', and that 'motivational' constructs could have a moderating effect on these assessments.

The results we have presented above would certainly seem to support the idea that cognitive and motivational constructs each have roles to play in forming teachers' overall opinions of pupils ability. By demonstrating that motivational constructs arise particularly from triads of Types B-F, our work would seem to suggest that teachers may be led to over- or under-estimate potential because of high or low motivation on the part of the pupil. We will return to this point later.

11.2 Preliminary conclusions in Physics

The over-riding conclusion from the work reported above is that physics teachers made use of fairly general concepts to do with ability in physics, and, of a wide range of motivational and personality based constructs.

It seems that they did not use constructs relating to the kinds of specific scientific behaviours that were mentioned on the checklists. This was even true of Teachers 03, 08 and 09 who had used the checklist in their identification.

One could suggest that the checklist had not made a lasting impression on teachers in terms of the way in which they looked at their pupils.

There is also an indication, mainly from Tables 11.4 and 11.5 that 'model pupil behaviour' (e.g. enthusiasm, confidence, conforming behaviour, concentration) acted as a favourite cue for teachers, and was associated with over-estimations of pupils' potential; and that qualities such as disorganisation and carelessness or quiet, withdrawn behaviour acted as disqualifying cues and were associated with under-estimation of potential.

11.3 Further analysis in Physics

Before presenting similar results in the other subjects we will pause to look a little more deeply at three matters concerning the physics constructs.

i) Our first concern is one that was raised in Chapter Ten: namely whether evidence can be found from the teachers' ratings of pupils to support the hypotheses that we have put forward that some constructs might have poles that were acting as favourite or disqualifying cues.

These hypotheses emerged from a consideration of the nature of the triads from which the constructs were elicited. If a teacher distinguished an under-estimated ($T\bar{P}$) pupil from two accurately judged (TP) pupils, one can certainly say that the pole of the resulting construct which described the $T\bar{P}$ pupil was associated with under-estimation of ability, and may have acted rather like a disqualifying cue, *for that individual pupil*. What cannot be said with certainty is that the teacher used this same cue in the same way in forming his/her judgement about *other* pupils.

If this could be established we could say that the cue is *consistently* associated with under-estimation on the part of the teacher - and that it would be worth attention as a *possible general cause* of under-estimation in the judgements made by that teacher. If it cannot be established, we can only make the weaker point that the cue was associated with under-estimation in the case of the individual pupil involved in the triad - and that it may be worth attention as something which may lead the teacher to errors of under-estimation on some occasions.

By looking at the way in which the teacher rated a number of pupils on the construct in question, and by comparing these ratings with the teacher's overall judgement of potential, we can begin to get some information that may help to clarify this point. We will discuss three examples, (a) - (c), below. In each case we will discuss pupils from the A or B Groups (i.e. pupils who had a test score which was high enough to imply that they were all good, or reasonable candidates for selection as more able). In making choices between these pupils, factors which affect the teachers' judgement are particularly likely to show up as the teacher is not in a position to respond simply to large differences in pupils' ability.

Example (a) Teacher 04H: a possible disqualifying cue.

In Table 11.5 'Slow to grasp ideas' was identified as a possible disqualifying cue for Teacher 04H. Table 11.10 shows how that teacher rated all of the A and B Group pupils whom he taught, on the construct containing that cue. We have divided the pupils into those whom he judged to be more able, and those whom he decided should not be in that category.

TABLE 11.10

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON CONSTRUCT 04H(b) FROM TABLE 11.5

	Slow to grasp ideas			Quick to grasp ideas	
	1	2	3	4	5
SELECTED	0	0	0	2	4
NON-SELECTED	2	1	2	4	1

Some comments are made on the statistical significance of the association between selection and rating on the construct in Note 8 at the end of this chapter. The note also comments, in similar fashion on Tables 11.11 to 11.13.

Although the numbers in this table are small, we see that 5 out of 5 pupils who were rated 1,2 or 3 (i.e. at or near the proposed 'disqualifying cue' pole) were denied membership of the top 10% group, despite the fact that these pupils, like all of the pupils in the table, had a reasonable test score. This confirms the idea that 'Slow to grasp ideas' may have acted generally as a disqualifying cue for that teacher. Interestingly a rating of 5 at the other end of the construct also seemed to distinguish between selected and non-selected pupils. The pole 'Quick to grasp ideas' seemed to operate as a favourite cue. Although on a priori grounds we have made no case that it should do so from the appearance of the construct in Table 11.5, the same construct does appear in Table 11.6 where we would wish to suggest that this role of favourite cue might be expected.

Example (b) Teacher 04H: a possible favourite cue.

Table 11.11 gives similar data in connection with a construct from Table 11.4 in which one pole, 'Confident in practical work', appeared to operate as a favourite cue.

TABLE 11.11

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON CONSTRUCT 04H FROM TABLE 11.4

	Hesitant in practical			Confident in practical	
	1	2	3	4	5
SELECTED	0	0	0	2	4
NON-SELECTED	1	1	6	1	1

37

In this example 6 out of 8 pupils who were rated 4 or 5 (i.e. at or near the proposed 'favourite' pole) were selected as more able by the teacher. It does seem that the pupil characteristic 'Confidence in Practical Work' was generally associated with teacher selection and may have been involved in tipping the balance in favour of a number of pupils when the teacher came to make up his list.

Example (c) Teacher 08: Possible disqualifying and favourite cues.

In Table 11.6 we suggested that the left hand poles may be favourite cues, and that the right hand poles of the same constructs may be disqualifying cues. Table 11.12 shows how Teacher 08 rated A and B Group pupils on one construct of this kind: 'Committed to the subject - not committed to the subject'.

TABLE 11.12

RATINGS OF SELECTED AND NON-SELECTED PUPILS ON CONSTRUCT 08 FROM TABLE 11.6

	Committed to subject			Not committed to subject	
	1	2	3	4	5
SELECTED	4	2	0	0	0
NON-SELECTED	0	1	3	1	1

6 out of 7 pupils rated 1 or 2 (at or near the proposed 'favourite' pole) were selected as more able by the teacher, and 5 out of 5 rated 3, 4 or 5 (at or near the 'disqualifying' pole) were not selected. This confirms the idea that 'Committment' operated as a favourite cue for this teacher in the assessment of the whole group of pupils, and that 'Non-Committment' operated, also in a general way, as a disqualifying cue.

Sometimes the tables needed to carry out these checks on a construct could not be constructed, either because very few A or B Group pupils were overlooked by the teacher (this was true of Teacher 03 and made it impossible to look at any of his constructs in this way), or because the range of ratings that was given to pupils was very small (this was true of the construct from Teacher 05 in Table 11.4). However the implication of these examples is that it is reasonable to suggest that the cues that were identified as disqualifying, or favourite cues on the basis of the nature of the triad from which they came, may well have operated in those same roles, in a consistent fashion, when the individual teacher went on to judge pupils who were not part of that triad.

38

All of the discussion above has related to further study of individual teachers. In the same vein it would be interesting, and potentially even more useful, to devise some further experimental test of whether the generalisations that we have made in discussing the nature of favourite and disqualifying cues (e.g. that 'model pupil behaviour' may act as a favourite cue for many of the teachers in our sample) are applicable to a large number of teachers. Such a study might be undertaken by asking teachers to judge whether pupils described in short pen sketches were able or not. The 'model pupil behaviour' element of the pen sketch would be varied systematically and an attempt could be made to see whether this variation had the effect that would be expected if this component of the description really did operate as a favourite cue for the whole group of teachers. We hope to be able to undertake such a study in the future, but can do no more at this stage than to suggest that it might be of value.

ii) We turn, now, to the second point that we wish to begin to explore here.

It may be useful to consider a classification of the constructs which have arisen in the study and to determine whether the constructs elicited from Type A triads (constructs associated with teachers' judgements which are in line with the pupils' test scores), those coming from triads of Types B to F (constructs associated with over- or under-estimation of potential) and those coming from random triads were of similar, or of different kinds.

Clearly such a classification should reflect the kinds of constructs which were elicited. However, in addition, there should also be some theoretical basis for each of the categories in the classification. Since we are interested in the relationship between teachers' constructs and pupils' future performance, we have worked on the basis that the classification should categorise constructs according to the kinds of factor which may be expected to influence pupil performance. Ability factors are clear candidates for inclusion. Youngman (1979) is one of many researchers who have shown that ability measures have considerable power as predictors of future performance. This has been confirmed in that part of our own work which is described in Appendix 1. In Chapters One and Two, and in Appendix 1, wehave discussed the role of specific and general aspects of ability. This discussion provides theoretical support for the inclusion of both specific ability and general ability components in the framework. The choice of physics ability and mathematics ability as the two sub sections of the specific ability component reflects the nature of the specific ability constructs that were elicited from teachers.

A considerable amount of research suggests that non-ability factors can also affect performance. Cohen and Manion (1981) point out that personality factors (especially extroversion/introversion) can have a considerable effect on performance. In addition, Entwistle (1968) established a clear relationship between motivation and achievement. It was therefore important to include these components in the framework. However, as colleagues who attempted to classify the constructs which we elicited, had some difficulty in agreeing whether some of these constructs were motivational or personality constructs, it was decided to combine these two categories in the final version of the framework that is set out in Figure 11.1.

Other influences on performance can also be traced in the literature. Brookover et al (1967) showed that pupil's academic self-image was correlated with grade point average. However it may be difficult for a teacher to be aware of a pupil's own opinion of his/her ability, and certainly no constructs relating to academic self-image were elicited. This component was therefore omitted from the framework.

Finally a wide range of situational factors such as teaching style (Eggleston et al, 1976), ability grouping (Postlethwaite and Denton, 1978) and general school ethos (Rutter, 1979) have been shown to relate to pupil performance to varying degree. However for a given teacher, in a given school, there was no variation in these factors across the pupils whom he/she was attempting to assess. As a result there were no constructs related to these factors, and they were therefore omitted from the classification framework.

Figure 11.1 summarises the classification which we designed in the light of these considerations.

FIGURE 11.1

A Framework for Construct Classification

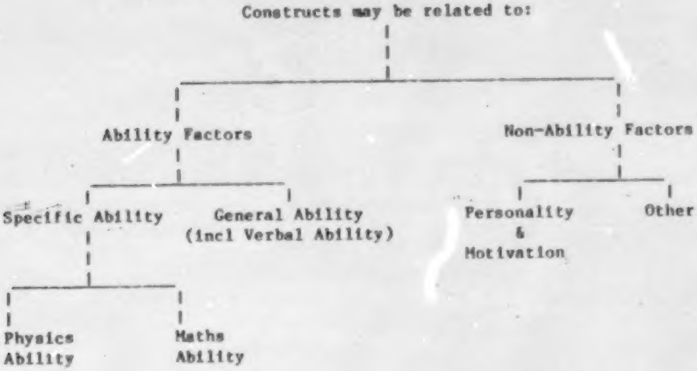

Constructs may be related to:
Ability Factors
Non-Ability Factors
Specific Ability
General Ability (incl Verbal Ability)
Personality & Motivation
Other
Physics Ability
Maths Ability

Table 11.13 shows the number of constructs of each type which fell into each part of the classification framework.

TABLE 11.13

THE NUMBER OF CONSTRUCTS OF EACH TYPE IN EACH CLASSIFICATION CATEGORY

Triad Type	Physics Ability	Maths. Ability	General Ability	Personality & Motivation	Other
A	4	2	4	3	0
B-F	6	1	14	32	2
Random	1	1	4	47	7

It is not surprising to see, from this table, that teachers were sensitive to factors such as the personality and motivation of their pupils. As indicated in the discussion above, these factors have a direct influence on the quality of the work done by pupils. However, such factors also colour the general relationship between the teacher and the pupil, and may be given too much weight when the teacher tries to predict the pupil's future performance.

It is interesting that the Type A triads which, on _a priori_ grounds were likely to give rise to constructs which were most likely to help the teacher to make judgements about pupils that were in line with the tests, were largely ability constructs. On the other hand, constructs from the triads of Types B to F (which we expect to be associated with judgements that were not in line with the tests) were largely personality, or motivational constructs.

The implication is that though motivational constructs may have some predictive power as indicators of future potential (a fact that teachers seemed to recognise as is indicated by the presence of some motivational constructs amongst those elicited from Type A triads), physics teachers did indeed seem to place too much weight on them in forming their overall view of pupils' potential; (as is indicated by the large number of motivational constructs amongst those elicited from Type B-F triads).

Teachers might therefore make judgements which more closely matched with the test results if they were to attempt to make a clearer distinction between their assessments of pupils on motivational/personality dimensions and their assessments on ability dimensions, and if, subsequently, they were to place less weight on the motivational/personality dimension in coming to their overall judgement. To this end the detailed nature of the motivational/personality constructs that we have noted in Tables 11.4 to 11.8 may be helpful in alerting teachers to the kinds of motivational/personality information which might be affecting their view of pupils. Of course, not only can we draw teachers' attention to this overemphasis on motivational constructs we can also, through the classification of characteristics as disqualifying or favourite ones, indicate the way in which these constructs were being misinterpreted. This detail may help teachers to modify their response to such characteristics in such a way as to bring their overall judgement of potential more closely in tune with the pupils' test results. It may be helpful at this point to remind the reader that a closer match between teacher and test judgement is a worthwhile goal in physics. As shown in Chapter Three, the extent of the mismatch between teacher- and test-based judgement is too great to be accounted for by test error alone. In bringing the teachers' judgement more closely in line with the test judgement we would argue that teachers would be improving their assessment of pupils' potential for future performance in the subject.

One further point arises from Table 11.13 and that is that general ability constructs were more apparent from triads of Types B to F than from Type A triads. By analogy with the argument above, this would suggest that teachers should try to distinguish more clearly between specific and general abilities and, subsequently, to place more weight on the specific ability judgements when forming their overall view of pupils' potential. This point is in complete accord with points made elsewhere in this report where overemphasis on general ability on the part of physics teachers has been noted.

iii) The third, and final area in which we would seek to make more detailed comment concerns the completeness of the construct data.

We mentioned in Chapter Ten that teachers may go about the process of forming an overall judgement of their pupils by combining the data which they have for a pupil on each individual construct by what amounts to a multiple regression procedure - albeit probably an unconscious one. We applied such a model to our construct data and generated, for each teacher, a combination of the individual construct scales that provided the best estimate of the teacher's overall judgement on his/her pupils. If we had a reasonably full map of the teacher's construct system, (and if teachers did actually use some procedure which was analagous to multiple regression), then we expect to be able to explain a fairly high percentage of the variance in the teacher's overall judgement by means of the composite score from the combination of constructs. Table 11.14 gives a measure of the variance which was actually explained in this way.

TABLE 11.14

THE MULTIPLE REGRESSION MODEL APPLIED TO TEACHERS' OVERALL JUDGEMENT OF POTENTIAL

Teacher	% Variance explained	no. of pupils	no of variables included in equation
03	72	27	3
04H	43	30	1
04M	31	33	1
05	45/23 (*)	28/28	1/1
07B	33	26	1
07G	23	26	1
08	62	28	2
09	40/61 (*)	22/24	1/1
10	19	29	1
11	60	33	1
12	57	37	1

(*) Teachers judgements on two separate classes were analysed separately.

In a large proportion of the cases in this table, a considerable amount of the variance in teachers' overall judgement could be explained on the basis of the construct ratings. For example in 61% of cases, 40% or more of the variance could be explained. There would therefore seem to be some evidence for the view that our map of the factors which influence teachers was complete enough to warrant the kind of detailed discussion which we have engaged in so far in this chapter.

11.4 Basic results in English

Tables 11.15 to 11.21 show the constructs which were elicited from English teachers. A short comment follows each table.

TABLE 11.15

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT TP pole	$\overline{\overline{\text{TP}}}$ pole
04A	Produces work of depth	- Does not produce work of depth.
04D(a)	Syntactically accurate	- Chaotic sentence construction.
04D(b)	Inventive writing	- Reliable, but just steady writing.
05C(a)	All round intellectual ability.	- Weak all round.
05C(b)	Exact vocabulary	- Imprecise vocabulary.
05R(a)	Not responsive	- Easy to deal with.
05R(b)	Clever, good background	- Less mature.
11A(a)	Generally able	- Generally weak.
11A(b)	Confident reading	- Hesitant reading.
11A(c)	Accurate work	- Inaccurate work
11H	Good O-level candidate	- Weak O-level candidate.

The main feature of these constructs (which were associated with teachers' judgements that were in line with the test results) was that they indicated sensitivity, on the part of the group of English teachers, to a number of apparently distinct English skills. This can be contrasted with comparable constructs in physics, where the only specific skill that was mentioned was that of mathematics ability. In common with physics teachers, English teachers also made reference to general ability in this table.

TABLE 11.16

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	CONSTRUCT		
	$\bar{T}P$ pole		$\overline{TP}$ pole
04A(a)	Confident	-	Not confident .
04A(b)	Not aware of what goes on in English.	-	Aware of what goes on in English.
04D	-		-
05C(a)	Strives for eloquence in emotional register.	-	Imprecise and inelegant.
05C(b)	Frequent signs of brilliance	-	Sound.
05R(a)	Thoughful and imaginative	-	Not evidently sensitive
05R(b)	Co-operative	-	Awkward in class.
11A	-		-
11H	-		-

We would remind the reader that the description at the left hand pole of these constructs may be acting as favourite cues for the teacher, leading him or her to nominate pupils for membership of the more able group who, in test terms, were unlikely to qualify for membership. In physics these favourite cues were largely to do with something that we labelled 'model pupil behaviour'. There are echoes of this in the English results (04A(a), 05R(b)). However one 'favourite cue' in English ('Strives for eloquence in emotional register') may be interpreted as an example of a teacher responding to an appropriate characteristic of high ability in English. This could have led to overestimation of the pupil in comparison with the test simply because the test was not able to measure such a sophisticated aspect of English ability. The same comment might apply to the favourite cue in construct 05R(a).

46

TABLE 11.17

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT TP̄ pole	 TP pole
04A(a)	Does not take work far enough	Takes trouble .
04A(b)	Not original in writing	Produces original writing
04D(a)	Unwilling to commit self to an opinion	Willing to commit self to an opinion
04D(b)	Does not contribute orally	Expresses opinions and asks questions
05C(a)	Shows moral sensitivity & is idealistic	Has accurate insights into the wider world
05C(b)	Carries own social world with him	Excited about social possibilities
05R(a)	Less forceful	Forceful
05R(b)	Not keen	Keen
11A	Poor vocabulary	Good vocabulary
11H	Poor vocabulary	Has a feel for words/Good

In two places in this table there are echoes of the kinds of attitudinal disqualifying cue that we found in physics (04A(a) and 05R(a)). However we can also see, again, evidence of a sensitivity, on the part of teachers, to subject-specific skills that may have been overlooked by the test (04A(b), 11A, 11H). We suggest that, under the TP̄ pole of these constructs, we find things which the test did not measure, and which the pupil did not do well. The teachers may have been influenced by these things in coming to their decisions that pupils were not more able. Since these subject specific weaknesses on the part of the pupils were not revealed by the test, the teacher and test disagreed. However the implication from the nature of these descriptions, is that it may be the test-based prediction which was in error.

TABLE 11.18

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT	
	$\bar{T}P$ pole	$T\bar{P}$ pole
04A(a)	Mature work	Immature work
04A(b)	Normally popular	Not popular.
04D	Good technical skills	Weak technical skills.
05C	Impassioned in use of language.	Steady in use of language.
05R(a)	Enthusiastic	Not forward.
05R(b)	More disorganised	Takes a lot of trouble.
11A	Gets comprehension exercises right	Gets comprehension exercises wrong.
11H	-	-

The constructs in this table seem to be consistent with those elicited in Tables 11.16 and 11.17. Most of the favourite cues (at the $\bar{T}P$ pole) and disqualifying cues (at the $T\bar{P}$ pole) that are found here can also be found, in identical or similar form, in the appropriate earlier table.

TABLE 11.19

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT	
	$T\bar{P}$ pole	$\bar{T}\bar{P}$ pole
04A(a)	Reliably good work	Sometimes produces good work.
04A(b)	Great sense of humour.	Reasonably quiet.
04D(a)	Good technical skills	Poor technical skills.
04D(b)	Assured sophisticated style	Underdeveloped style.
05C	Does not miss the point	Intellectually weak.
05R	Confident	Lacks self confidence.
11A	Has natural ability	Has no natural ability.
11H(a)	Generally able (written work grammar, sentence construction)	Not able in these ways.
11H(b)	Good O-level candidate	Weak O-level candidate.

If one attempts to interpret the left hand pole of these constructs as disqualifying cues, one is led into some difficulty as most of them refer to high subject-specific ability. In this respect English and physics teachers

48

differed, as most of the corresponding physics constructs (Table 11.7) mentioned poor attitudinal characteristics which did fit sensibly into the disqualifying cue model. The apparent inconsistency in English can be resolved by referring back to Table 10.1 where we first set out the interpretation of constructs arising from critical triads. In connection with Type E triads, we put forward an alternative interpretation. This was that the teacher may have placed too little weight, on the characteristics described under the $\overline{\mathrm{TP}}$ pole, as indicators of high potential in the subject. From the nature of the constructs, this would seem to be the more appropriate interpretation in English.

TABLE 11.20

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT		
	$\overline{\mathrm{TP}}$ pole		TP pole
04A	Not aware of what is happening in lessons	-	Conscientious and studious.
04D(a)	Unimaginative	-	Imaginative.
04D(b)	Not committed to the subject	-	Committed to the subject.
05C	Satirical streak	-	Makes conventions work.
05R	Clever but quiet	-	Clever and not quiet.
11A	-		-
11H	Essays not sensitive	-	Sensitive Essays.

The $\overline{\mathrm{TP}}$ pole of Construct 05C seems to be one which could be interpretted as a favourite cue. Other constructs (espcially 04D(a) and 11H) would seem to be things to which the teachers gave too little weight as indicaters of low potential for the subject.

It may seem to be inconsistent to argue that the constructs arising in connection with triads of type B-F were, in physics, indicative of pupil characteristics which led teachers into errors of judgement, whereas, in English, they were largely indicative of appropriate characteristics which the teachers noticed, but which the tests did not measure. However, we would suggest that two things support this difference in interpretation. First, it must be remembered that, in Chapter Three, we found evidence that physics teachers did make errors of judgement in assessing pupils' potential for O-level, whereas we found no evidence to suggest that English teachers made such errors. It was possible to account for the full extent of the mismatch between English teachers judgements and English test scores entirely in terms of error in the *test*. Secondly the nature of the constructs themselves were consistent with the two different interpretations. Indeed one might argue that the consistency between the findings of Chapter Three, and the nature of the kinds of constructs found in English and physics was indirect evidence, suggestive of the validity of the critical triad methodology.

TABLE 11.21

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT *A PRIORI* CLASSIFICATION

TEACHER	CONSTRUCT	
04A(a)	Lively, enjoys contributing orally	- Fairly passive in class
04A(b)	Mature	- Immature
04A(c)	Nervous, anxious to do well	- Confident
04A(d)	Generally more able	- Generally less able
04A(e)	Depth in work	- Superficial work
04A(f)	Unable to relate to what is going on	- Integrated in to work of class
04A(g)	Poor written work technically	- Technically competent in written work
04D(a)	Careful fastidious worker	- Driven by inspiration
04D(b)	Mature - grown up for age	- Not grown up for age
04D(c)	Makes witty remarks	- Makes few remarks
04D(d)	Does not read books	- Usually has a book on the go
04D(e)	Lazy/slipshod	- Works with care and accuracy
04D(f)	Tries to incorporate new into work	- Does not incorporate new ideas
04D(g)	Little improvement in quality of work	- Improvement in quality of work
04D(h)	Sullen	- Cheerful
04D(i)	Not good at listening to instructions or advice	- Takes note of instructions and advice
05R(a)	Articulate	- Reserved
05R(b)	Noticeable personality	- Retiring personality
05R(c)	Confident	- Quiet
05R(d)	Gives in class	- Does not contribute in class
05R(e)	Chatty in class	- Restrained in class
05R(f)	Responsive to English	- Amenable but not responsive
05R(g)	Diffident	- Not diffident
05R(h)	Supportive to teacher	- Regular member of class
05R(i)	Does not take part in class	- Part of lessons

Table 11.21 is continued.

Table 11.21 continued

TEACHER	CONSTRUCT	
05C(a)	'Meaty' use of language	Stylistic use of language
05C(b)	Little talent	Talented orally/practically
05C(c)	Poor stamina for work	Has stamina
05C(d)	Exhibits deftness and control	English is an uphill struggle
05C(e)	Perceptive, imaginative	Not receptive or imaginative
05C(f)	Doesn't understand English tasks	Does understand English tasks
05C(g)	Reliably sound	Assertively bright/sensitive
05C(h)	Sober seriousness	Not guided by predetermined standards
05C(i)	Light	Strives for depth
10T (a)	Quiet	Fluent orally
10T(b)	Excellent reading aloud	Poor reading aloud
10T(c)	Consistent top stream performance	Late developer
10T(d)	Quick thinking - work comes easily	Ponderous worker
10T(e)	Idle	Hardworking
10T(f)	Needs to be drawn	contributes without being drawn
10T(g)	Outstanding vocabulary	Limited vocabulary
10T(h)	Excellent oral work	Little oral work
10T(i)	Mechanically correct	Makes mechanical errors
10T(j)	Flair in written work	Mundane written work
10T(k)	Lacks confidence in work	Confident
10T(l)	Not a reader	A great reader
10T(m)	Sensitive artistic expression	Ordinary expression
10T(n)	Accurate description	Inaccurate description
10T(o)	Highly articulate	Not articulate
10T(p)	Writes beautiful poetry	Writes laboured poetry
10T(q)	Concise clear expression	Expression not concise and clear

Table 11.21 is continued.

Table 11.21 continued.

TEACHER	CONSTRUCT	
11H(a)	Fluent writing - more ideas	Not fluent
11H(b)	Sensitive, expressive writing	Does not write sensitively
11H(c)	Poor comprehension	Good comprehension
11H(d)	Makes basic mistakes	Accurate
11H(e)	Good vocabulary	Poor vocabulary
11H(f)	Good grammar	Inaccurate grammar
11H(g)	Good ideas for essays	Weak ideas for essays
11H(h)	Lively mind	Not a lively mind
11A(a)	No interest	Interested
11A(b)	Poor attention	Attentive
11A(c)	Poor written work	Good written work
11A(d)	High accuracy in spelling and punctuation	Inaccurate
11A(e)	Tries hard	Does not try hard
11A(f)	Takes active part in lessons	Not involved in lessons
11A(g)	Lacks maturity of thought	Mature thinker
11A(h)	Gives up with deeper work	Perseveres
11A(i)	Generally able	Lacks ability
11A(j)	Has flair for use of words	Shows no flair
11A(k)	Has interesting style	Ordinary style

* This teacher's nominations matched test nominations so closely that no critical triads could be constructed.

11.5 General conclusions for English.

The main conclusions to be drawn from the English work are to be found in the brief discussions following each table of constructs. However it is interesting to offer some general, concluding remarks. We note that some 64% of English teachers' constructs were concerned with pupils' ability (48% being concerned with specific English ability). This can be compared with physics where only 29% of teachers' constructs were related to pupils' ability (and only 11% were related to specific science ability). In view of the greater accuracy of English teachers in identifying more able pupils, this is an important general point.

A subjective view of English teachers also emerged in the course of the elicitation interviews. We found that they were quick to respond with a considerable amount of information about individual pupils, and that this was often expressed with enthusiasm and genuine concern. For example, the teacher who described a pupil as writing 'beautiful' poetry really did mean the word beautiful and had incorporated some of this pupil's work into her own collection of poetry to be read for pleasure. It is encouraging that such involvement in the work of pupils is possible and interesting that this group of teachers proved to be so successful in their overall judgements.

11.6 Basic findings in French.

We now present the results of the study of the constructs used by French teachers. Unfortunately it was only possible to interview four teachers of this subject so the insights gained cannot be as full as they have been in physics and English. The main value of the work in French will be in indicating the kinds of constructs that teachers use. Generalisations about the nature of favourite and disqualifying cues will be difficult because of the small numbers of constructs of each type that can be studied. Nevertheless we will present the data in the same form so that it is easy for the reader to make some broad comparisons between the subjects should he wish to do so.

We give first the constructs arising from Type A triads.

TABLE 11.22

CONSTRUCTS FROM TYPE A (A1 and A2) TRIADS

TEACHER	CONSTRUCT	
	TP pole	$\overline{TP}$ pole
03	Highly able	General lack of ability.
04J	-	-
04T(a)	Outstanding orally	Lacks oral confidence
04T(b)	Generally competent	Low general ability in French
05	Confident	Not sure of self

It is interesting to see confidence constructs (05 and 04T(a)) in this table, (which, we have argued, is the table containing constructs which are appropriate to the judgement of potential in the subject), and to see the close link that is implied, in construct 04T(a), between confidence and ability. The results seems to some extent to be at odds with English and physics where the corresponding tables contained mainly ability constructs. The finding is, however, entirely in line with Pimsleur's (1964) view which we have quoted in Chapter Five. He stated that:

> ...the person who does well in a foreign language is one who is willing to get up and talk before others...

Other subject specific skills were not mentioned (in contrast to English where such skills were a major feature of the table). This does suggest that the English teachers' ability to notice these skills was not simply the result of the regular daily contact with particular groups of pupils (for the French teacher also enjoys this advantage in most schools). It argues that it is something in the nature of English teaching which enables the teachers to be aware of specific abilities in that subject.

In Tables 11.23 to 11.27 we give the constructs from the triads of Types B to F. In view of the very small numbers of constructs in each table, we will resist the temptation to draw further conclusions from each one.

TABLE 11.23

CONSTRUCTS FROM TYPE B TRIADS

TEACHER	CONSTRUCT $\bar{T}P$ pole	 $\overline{TP}$ pole
03	-	-
04J	Generally more able in French-	Less able in French
04T	-	-
05	Sees patterns in French - clearly	Does not see patterns clearly.

TABLE 11.24

CONSTRUCTS FROM TYPE C TRIADS

TEACHER	CONSTRUCT T$\overline{P}$ pole	TP pole
03	Competent	Outstanding linguist
04J	Lacks imagination	Has imagination
04T(a)	Underperforms	High quality performance in subject
04T(b)	Poorly motivated to subject	Well motivated to subject
05(a)	Extrovert	Quiet academic
05(b)	Careful/fastidious	Slapdash

TABLE 11.25

CONSTRUCTS FROM TYPE D TRIADS

TEACHER	CONSTRUCT $\overline{T}$P pole	T$\overline{P}$ pole
03	-	-
04J	Uses talent	Under-uses talent
04T	-	-
05(a)	Able	Not able or motivated
05(b)	Well motivated	Little committment

TABLE 11.26

CONSTRUCTS FROM TYPE E TRIADS

TEACHER	CONSTRUCT $T\bar{P}$ pole	$\overline{TP}$ pole
03	Serious about work in French	Underachieving
04J	-	-
04T	-	-
05	Brighter, more motivated	Afraid to let herself go

TABLE 11.27

CONSTRUCTS FROM TYPE F TRIADS

TEACHER	CONSTRUCT $\bar{T}P$ pole	TP pole
03	-	-
04J	Has flair/natural ease with subject	Systematic/hardworking
04T	-	-
05(a)	Determined to make progress	Satisfied with own level of achievement
05(b)	Good at applying knowledge	Weak in applying knowledge

Finally, in Table 11.28 we give the constructs arising from random triads, and from critical triads that were not split in the expected way.

TABLE 11.28

REMAINING CONSTRUCTS ARISING FROM TRIADS WHICH DID NOT PERMIT A PRIORI CLASSIFICATION

TEACHER	CONSTRUCT		
03 (a)	Highly competent	-	Has difficulties with technical details
03 (b)	Bad attitude to the subject	-	Enthusiastic for subject
03 (c)	Irrepressible gossip; talks not works	-	Able to concentrate
03 (d)	Quiet serious worker who achieves	-	Achieves little
03 (e)	Over pressured	-	At ease with work in French
03 (f)	Has confidence to participate in oral work	-	Lacks oral confidence
04J(a)	Lacks application	-	Has application
04J(b)	Unreceptive of advice from teacher	-	Keen to please
04J(c)	Gives impression of brightness	-	Steady
04J(d)	Alert	-	Vague
04J(e)	Impulsive	-	Painstaking
04J(f)	Seeks attention	-	Reclusive
04J(g)	Independant of mind	-	Amenable
04J(h)	Orally confident (extrovert)	-	No oral confidence (introvert)
04J(i)	Easily hurt	-	Resilient
04J(j)	Insouciant	-	Concerned to please
04T(a)	Confident	-	Not confident
04T(b)	Extremely quiet; gives little orally (or written)	-	Confident and lively
04T(c)	Willingly contributes	-	Contributes only if asked
04T(d)	Scruffy written work	-	Meticulous
04T(e)	Academically outstanding	-	Sloppy
04T(f)	High academic standard	-	Low academic standard
04T(g)	Extremely quiet	-	Positive
04T(h)	Sociable	-	Restrained
05 (a)	Highly able in languages	-	Lacks sense of purpose & effort
05 (b)	Co-operative	-	Awkward

11.7 General conclusions in French.

Confidence (particularly in relation to oral work) and motivation were clearly important components in the French teachers' maps of childrens' abilities. There were also several mentions of general 'Ability in French'. However, there were few references to subject specific skills, such as technical accuracy, memory or vocabulary, which we might have expected.

Some 35% of constructs were clear ability constructs, 10% related to confidence, 20% to motivation and 35% to clear non-ability constructs. The comparison between English and French teachers in terms of the kinds of constructs which they used is again of interest in view of the generally lower accuracy of French teachers in identifying the more able.

Perhaps the most important general conclusion is that, because of the small number of teachers who were involved, the work which is reported here should be considered as a pilot study of French teachers' construct systems. A larger follow up study might add much that is of interest - including insights into the nature of favourite and disqualifying cues used in this subject.

Constructs used in Mathematics.

As mentioned at the beginning of this chapter, a simpler study of mathematics teachers' constructs is provided in Appendix 8.

Notes for Chapter Eleven

1) It is interesting to note that one particular teacher (08) was responsible for four of the 12 unsplit critical triads.

2) Teacher 05 was no longer teaching in the same school when the interview was conducted and one would expect impressions of past pupils to be somewhat overlaid by new concerns. Teacher 11 was interviewed in a rather noisy staffroom.

3) F1 triads were the first type of critical triad to be presented and many teachers found two ways of splitting these triads. The splits for this triad type were not therefore independent, and this particular significance figure should be treated with caution.

4) The chance of making 8 expected choices from the 12 possible choices if the selection was based on a random split of the triad is

$$^{12}C_8 \cdot (\tfrac{1}{3})^8 \cdot (\tfrac{2}{3})^4$$

The significance level = $\sum_{n=8}^{12} {}^{12}C_n \cdot (\frac{1}{3})^n \cdot (\frac{2}{3})^{(12 - n)}$

5) The significance level associated with 6 expected choices out of 11 is 0.12 but that associated with 7 out of 11 is 0.04.

6) It may be helpful to remind the reader that the teachers did not know the pupils' test scores when they made their overall judgement of the pupils' potential, nor when they took part in the elicitation interview.

7) Two constructs from this table (04H and 07G) call for further comment as the descriptions found at the right hand poles of these constructs seem difficult to reconcile with the fact that the pupils who were being described were high scoring pupils who had been nominated as more able by their teacher (TP pupils). One is tempted to ask how it can be that a TP pupils can be described as 'Slow to grasp ideas' or as pupils who achieved only 'Low performance' levels. The solution to this problem in the case of Construct 04H is fairly straightforward. Within the triad, the two TP pupils were described as 'Slow to grasp ideas', in contrast

to the $\overline{\text{TP}}$ pupil who was described as 'Quick to grasp ideas'. However when the teacher rated all of his pupils on this construct, all three who had formed the triad were rated as 'Quick'. One can interpret this as indicating that the $\overline{\text{TP}}$ pupil was particularly quick, even in comparison with the two TP pupils, but that all three were quick in comparison with the rest of their classmates. This interpretation removes the puzzle over the TP pole, without calling into question the possibility that 'Quick to grasp ideas' was operating as a favourite one for that teacher.

Similar problems of this magnitude do not seem to have arisen in connection with most of the other constructs reported in this section.

We turn now to the apparent difficulty concerning the interpretation of the right hand pole of Construct 07G. One of the TP pupils who was described as showing 'low performance' can be accounted for by an argument similar to that set out above. His rating on the construct indicated that his performance was perceived as being lower than that of the $\overline{\text{TP}}$ pupil in the triad, but it was still high compared with the rest of his class. For the other TP pupil in the triad the explanation of the apparent inconsistency is more complex. As can be seen from Construct 07G(a) in Table 11.9, this teacher attempted to distinguish between 'ability' and 'performance'. The second TP pupil who was described as showing low performance when the Construct 07G in Table 11.8 was elicited, was rated in such a way on Construct 07G(a) (Table 11.9) as to suggest that his ability was perceived as being higher than his level of performance. This higher ability on the part of the pupil was entirely consistent with his classification as TP.

8) The numbers in this table were too small to allow a chi-squared test to be carried out on the 5 by 2 table. Instead the table was regrouped as a 2 by 2 table (ratings 1 to 3 were combined, as were ratings 4 & 5). A Fisher Exact Probability test confirms that the new table showed a significant association between selection and rating ($p=0.05$). An identical re-grouping of Table 11.11 was carried out and again a significant association was detected ($p=0.005$). Table 11.12 was re-grouped by combining ratings of 1 & 2 and ratings of 3 - 5. This again showed a significant association ($p=0.01$).

Table 11.13 was re-grouped into ability and non-ability factors to overcome small number problems. The resulting table showed a significant association between the triad type and the classification of the resulting construct ($X^2=27.3$, $df=2$, $p\leq0.001$). As can be seen from the table below the Type A triads gave rise to more ability constructs than expected, whereas the random triads gave rise to non-ability constructs.

Triad	Construct Classification	
Type	Ability	Non-Ability
A	10 (3.75)	3 (9.25)
B-F	21 (15.89)	34 (39.11)
Random	6 (17.34)	54 (42.66)

Observed numbers are shown in the table with the relevant figures from the expected table in brackets.

If Type A triads were compared just with Type B-F triads, a significant difference between the two was found ($X^2= 4.89$, $p\leq0.05$). Again the Type A triads gave rise to more ability constructs and the Type B-F triads to more non-ability constructs.

CHAPTER TWELVE

Further Development and use of Checklists

12.1 Introduction

From the preliminary work with checklists it is clear that at the very least a different method of use must be employed if there is to be worthwhile effect on the identification of children of high ability in the classroom. The level of knowledge that a teacher has of each individual would seem to be insufficient for an adequate profile to be built up when the lists are filled in at one sitting. In other words it is not sufficient to merely draw the teacher's attention towards the appropriate characteristics for ability in the subject and away, say, from a markbook of accumulated evidence of performance, which gives at best a general overview of achievement over a period of time.

This is disturbing in itself because it shows that teachers may not have this important knowledge available. Even though, in such subjects as English and mathematics and less so in such subjects as physics and French, the general knowledge of subject specific ability may be adequate to predict within acceptable limits for success of O-level, the fuller knowledge of the profile of aptitudes is important if provision is to be tailored to the needs of the individual.

Furthermore, if a teacher cannot rely on his/her ability to gain detailed knowledge from day-to-day interaction with the pupils, he/she becomes underconfident in this ongoing method of monitoring progress, so that the benefits for the individual disappear. These benefits can be compared with the consequences of the once and for all method of administering a battery of psychological tests and the resulting labelling of children.

The conclusions that we have drawn indicate a problem, but we do not intend to make unjust accusations. The secondary teacher works under considerable constraints such as time, class size and syllabus, so that it is within these constraints that the identification process must proceed.

It may be possible to overcome these constraints to some extent by diverting more of the limited energies available, for the numerous tasks in the classroom, to attempt to improve the detailed identification guided by the checklists. It was with this in mind that we sought the co-operation of a sample of teachers in each subject to use the checklists in an appropriate way on new groups of third form pupils, during the second year of our research.

There were signs, from the first year, that the task was not completely hopeless, and that what was required were refinements to the method rather than dramatic change. The following analysis gave us some encouragement in this respect.

When we distributed the checklists for the first round, we asked teachers to fill them in for the top 20% of the year group and then to identify the top half of this group taking account pf checklist results. We also asked them to fill in the checklists for randomly selected pupils whose predicted O-level scores put them below the 80th percentile. In this way we could establish whether, even if the checklists could not separate abilities in the higher bands, they may still have enabled coarser identifications to be made. We scrutinised the completed checklists to try to discover if, in general, pupils ranked lower by tests were scored lower on the checklists. We would not wish to say, at this stage, that tables like the one below do any more than give us some encouragement that the broad separation of abilities is possible. We have not attempted to extract more from this data than is justifiable in the light of this point. For example we have not subjected the data to statistical tests. Nevertheless, we feel that there is some evidence that, in general, very weak pupils were scored lower than very good pupils. Table 12.1 gives figures for one school and subject. The distribution of percentages that is implied from this table illustrates what was to be found for each school and subject. The figures were derived from the number of checklist ticks in the Low, Medium and High categories, totalled for all checklist items. The results are given separately for pupils in each of the three bands: top 10% on checklist aided identification, next 10%, and random (below 80th percentile on the tests), and are expressed as percentages of the total number of checklist ticks.

TABLE 12.1

PERCENTAGE NOMINATIONS AS HIGH, MEDIUM OR LOW ABILITY ON CHECKLIST ITEMS, TOTALLED FOR ALL PUPILS (SCHOOL 10,MATHS)

	LOW	MEDIUM	HIGH
TOP 10% on checklist aided identification	0%	14%	86%
Next 10%	4%	56%	40%
Random below 80th percentile	57%	28%	15%

It may be that we are wrong to interpret these results as showing that checklists could, at least, be used for coarse separation of abilities. It could be that teachers have already labelled pupils into ability categories and merely fill in the checklists to reflect this labelling. This possibility is quite plausable if the detailed knowledge of the spectrum of abilities is not present even in the coarser sense. Nevertheless we lose nothing by treating the results optimistically and proceeding to our second phase of activity.

In the previous two chapters we have shown how a certain number of our sample of schools contributed to the other aspect of our study. This left a further restricted sample to co-operate with the development of checklists. The lack of difference between all schools in the first years work allowed us to work with these limited samples to discover general results that would apply to the wider sample.

The general strategy in English, French and physics was to negotiate the manner of checklist usage with each teacher, first to establish the relevant checklist items for the teacher's own broad aims and second to agree on the most appropriate method of use for the particular school environment. We were able to choose just one subject, because of manpower limitations, for more intensive classroom based activity. In all subjects, however, the general strategy was for the teachers to have the checklists at hand and in mind during the planning and teaching of lessons, so that the clues displayed, from time to time by the pupils, could be recorded and accumulated for a final scrutiny when it was considered that sufficient recording had taken place for valid identification of the top 10% to be undertaken.

Appendix 5 includes details of the developed checklists and recording forms.

We will now summarise the main points regarding checklist development for English, physics and French and then report on the analysis of results for these three subjects. We will then give a fuller report of the work done in mathematics before drawing final conclusions together.

12.2 Development of further use of checklists in English, French and Physics

ENGLISH

Discussions with teachers that used Year 1 checklists revealed no strong or consistent objections to individual items. In nearly every case, a small proportion of items (one or two) was identified as less satisfactory than the majority of items. Because of the fact that different English teachers, with their different approaches to the subject, invariably picked out different items for criticism suggested that no single item could be identified as inadequate. All items gained enough support to justify their inclusion in the second round checklists.

The chief lesson from Year One was the need for both a longer period of time in which to apply the checklists and a more simple and concise format. Thus, in conjunction with the other subjects, the English checklists were applied over one and a half terms. The format was redesigned in such a way as to eliminate the continuum approach, and a single description of each item was provided.

Each head of department who agreed to use the checklists was consulted over both the method of applying the lists, and over the choice of items. Where dissatisfaction over particular items was expressed, the teachers who would be using the lists were encouraged to modify them to suit their own criteria or means of assessment.

In every case, the departments involved entered into the process assured of the flexibility available to them in applying the lists, and the teachers involved appeared satisfied with the requirements of the task they had undertaken.

It transpired, as the term and a half elapsed, that the increased period of time did not - in most cases - lead to an increase in teachers' ability to use lists with confidence. Teachers who at the start had every intention of carefully checking the items over a long period of time still found it easier to deal with them in a more intensive way at some stage towards the end of the process. Almost without exception, they attributed this to the fact that they were being asked to apply criteria that - even if similar to those they might themselves select - had not been generated by themselves or (most crucially) by the work they had undertaken in that time.

This effect was evident consistently in all the schools (08, 09, 04, 00) using the checklists. The lists were filled-in in all but one case (half of school 09) but with little sense of satisfaction or accuracy in the opinion of the teachers. Most felt that previous knowledge and assessments of the individual pupils was applied onto the lists, rather than that the lists elicited new information about the pupils. Teachers felt that the exercise was artificial, in this form, although many did feel that the idea of a checklist was, in itself, valuable if it could be developed out of the syllabus that they had devised in their own department.

FRENCH

Although all the French teachers who used the checklists in Year One did succeed in completing them, there was a strong sense of dissatisfaction expressed over what they felt was a hurried and ineffective procedure. The short time-span seemed to accentuate problems with the items on the checklists: a lack of familiarity in the working of certain items emphasised the rather alien nature of this task. Unfortunately, as became apparent during discussion with teachers at this time, their existed very little middle ground between the kinds of item that elicit no new kind of insight about pupils (familiar concepts in familiar terminology) and the kinds of item that were meant to direct teachers to new insights about their pupils (unfamiliar concepts, or unfamiliar terminology) but turned out, in practice, to be more baffling than revealing.

The most evident fact at this stage was the limited number of criteria which French teachers felt were valid in assessing high ability in third years pupils. A need was clearly felt for a limited number of familiar and therefore clear items - but it was hard to see how a checklist designed along such lines could elicit fresh insights.

New checklists, with the least approved items discarded and certain new suggestions incorporated, were offered to teachers for Round Two. Teachers were encouraged to modify any individual items that they found excessively unsuitable, and were also encouraged to complete the lists steadily over the full one and a half terms.

Discussions with the four departments involved produced a fair consensus of approval for both items and the method proposed.

At two of the schools (04 and 08) the lists were used as planned and the items were greeted with a fair amount of approval, with the normal reservation that lists generated by the teachers themselves would be more effective.

At another two schools (09 and 00) the increased amount of time produced minimal success in achieving successfully completed lists. In both departments, the explanation was offered that the items did not relate, in practice, to the kind of work that was being carried out during this period.

Although the idea of the checklists met with general approval, it was obvious that there were many difficulties involved in using lists that were designed externally. One teacher summed up some of the difficulties as follows:

> One thing that comes out of the provision of the checklists is that we do not seem to have made very extensive use of them! Our intentions were no doubt always good, but unless we made a conscious effort at the time to take them out in class I certainly found that they were left incomplete for long stretches, and cannot be regarded as reliable indicators of anything more than observations made at irregular intervals.

PHYSICS

After they had completed the first version of the physics checklists, the teachers involved were consulted, so that ideas for the improvement of the lists could be accumulated. It transpired, from these discussions, that teachers were generally happy with the items on the checklist (though there were a very small number of instances where teachers were unsure about the exact meaning of individual items). However several aspects of the use of the lists caused problems for many of the teachers.

First teachers felt that it was very difficult to fill in the list reliably at one sitting. Some aspects of the pupils' scientific behaviour which the lists required teachers to assess, were only clearly displayed in connection with particular parts of the syllabus (indeed this was recognised in the design of the list by indicating syllabus areas where the teacher might expect to be able to observe the characteristics specified by the items.) Teachers felt very unsure of their ability to recall, sometimes after quite long periods of time, how each pupil matched up to each of the separate descriptions of scientific behaviour set out on the list. Several teachers mentioned that they were aware of the possibility of being influenced by a halo effect; so that a pupil known to perform in the way described by one item, would be assumed to be equally competent in terms of the other items.

Secondly teachers commented on a futher consequence of the fact that the opportunities for pupils to display some of the behaviours set out on the checklist were fairly rare. They felt that it was difficult to give a true picture of a pupil when the checklist asked them to comment on whether that pupil 'always, often rarely or never' behaved in accordance with the description of behaviour that was given on the checklist.

Finally some teachers felt uneasy about their ability to interpret the checklists once they had been completed.

Since there had been a little lack of clarity in the items themselves we decided, in preparation for the use of modified checklists in the second year of the project, to discuss each item with each of the teachers who were actually going to use the checklists. (As we pointed out in Chapter Five discussions with teachers had been an essential part of the design of the lists - but not all of the teachers who filled in checklists had been involved in those discussions.) These discussions were followed up by a summarising note and an example of such a note is included in Appendix 7. These discussions certainly clarified the thinking behind the items and also gave teachers the opportunity to reject individual items as being irrelevant to the particular kind of course which they teach in the third year. Teachers took very little advantage, in practice, of this opportunity to veto items.

To overcome the problems of completing the list at one sitting and of responding appropriately to the request to say whether pupils 'often, usually, rarely or never' displayed the high level behaviours indicated on the list we changed the format to that indicated below:

2nd YEAR PHYSICS CHECKLIST FORMAT

	PUPILS NAMES					
ITEM	Ann	Sarah	....	John	....	Simon
Selective in observing						
Seeks to explain not just describe phenomena						
....						

A full version of such a checklist grid is included in Appendix 7. Teachers were asked to use the checklist over a period of 1½-2 terms and to indicate, with a tick in the appropriate box, in a given child displayed one of the behaviours listed.

Although it would have been tempting to solve the problem of intepretation by asking teachers to count the number of ticks which accumulated for each child and to give, as their top 10%, those children with the most ticks we were still sceptical of the wisdom of using the lists in such an automatic way. As mentioned in Chapter Five, items varied in importance and, anyway, the checklist was intended to be used alongside other information which the teacher accumulated on his/her pupils. We resolved this problem by agreeing that the final selection of the top 10% pupils should be done by the teacher in consultation with one of the research officers. The procedure followed in most of the schools was that the teacher made tentative nominations for the top 10% list. The teacher and research officer then consulted the checklist and discussed the nominations taking particular care over pupils who seemed to be in the list of nominations but not to have many ticks on the checklist, and those who were not on the list of nominations despite a considerable number of ticks.

The teacher then made the final decisions and produced the top 10% list.

This procedure for completing and interpreting the lists was much more acceptable to the teachers - although, without exception, they did find it difficult to observe and record the scientific behaviour of their pupils in the midst of all of the other tasks which have to be undertaken during, and immediately after lessons. For some items few, if any ticks were recorded. This must throw into doubt the validity of the record provided by the fully completed checklists that were returned during the first year of the project, and may be one reason (perhaps the main reason) for the apparent ineffectiveness of the lists in improving teacher identification procedures on that occasion.

Effectiveness of the checklists

Before turning our attention to the more detailed work that was possible in mathematics we will look briefly at quantitative aspects of the second identification by checklists. The most appropriate table for this purpose is Table 12.2 which is rather like others from previous chapters, where we counted up how many identified pupils were in the categories 'good', 'satisfactory' and 'unlikely', based on their predicted O-level score. The reader may wish to refer to Section 3.4 to be reminded of the background.

We have compared the distribution of numbers of pupils who were nominated in these categories with the distribution of numbers identified from the first year's work without checklists. The comparison was by means of a one row chi-squared 'goodness of fit' test and was a way of establishing any quantitative change in the distribution of pupils over the three categories due to checklist usage.

The figures from the first year's work are extracted from Table 3.10.

TABLE 12.2

COMPARISON OF THE DISTRIBUTION OF PUPILS WHO WERE NOMINATED AS IN THE TOP 10% BY THEIR TEACHERS WHEN CHECKLISTS WERE USED WITH THE DISTRIBUTION WHEN CHECKLISTS WERE NOT USED

		NO. OF CHECKLIST AIDED NOMINATIONS	% OF NOMINATIONS WITHOUT CHECKLIST AID	EXPECTED NUMBERS FOR GOODNESS OF FIT TEST	X^2	SIGNIFICANCE
ENG	GOOD	35	61	31.11	3.48	NS
	SATISFACTORY	8	27	13.77		
	UNLIKELY	8	12	6.12		
FR	GOOD	23	51	24.48	1.41	NS
	SATISFACTORY	9	13	6.24		
	UNLIKELY	16	36	17.28		
PHYS	GOOD	26	45	18	6.54	5%
	SATISFACTORY	7	25	10		
	UNLIKELY	7	30	12		

We see that in English there was not a significantly different distribution of pupils in the three categories. We recall that in English there was no major concern from the preliminary work about the distribution of pupils in the three categories in schools that did not use checklists. The concern does still exist, however, that the teachers found difficulty in using the checklists so that their identifications via checklist were based, perhaps, on the same sort of knowledge as without them. So, although in English, their unguided assessments of pupils seemed very appropriate, the credit for this cannot, it seems, be attributed to checklist use.

Neither round of checklist usage has increased the level of French teachers' success in identifying the more able pupil. There is therefore the same indication, as in English, that the teachers were operating on the same kind of information whether they completed checklists or not. Of course the added problem in French is that earlier work, reported in Chapters Three and Eleven, suggested that this information did not seem to lead to accurate identification, nor did the component parts of it seem to be very appropriate.

In physics we see that significantly more pupils were chosen in the 'Good' category from this round, so that we are encouraged to believe that checklists can lead to improvement in the identification strategy, but we would not underestimate the effort that is required from the teacher who wishes to discover appropriate knowledge of the pupils' strengths and weaknesses in this way.

We can add to the above results by calculating the percentage of the high scoring A Group pupils who were identified from this checklist aided activity.

The percentages were: English 89%, French 61% and physics 83%. There would seem to be some concern in both physics and English that a number of these pupils were still missed, but a major concern in French that as many as 39% were missed.

12.3 Mathematics

Introduction

We will now describe the more detailed work that was possible in mathematics, starting with a description of the research aims and methodology.

One reason that checklists were not a very effective tool in the identification procedure could have been that the sorts of clues that are available to teachers both in the classroom and from scrutiny of pupils' workbooks, are inadequate. This could be associated with the inadequate challenges presented to the pupils in their work, so that the responses to these challenges would inevitably give the teacher inadequate information, or it could be that the environment in which the clues to ability occur will not allow adequate observations and records to be made.

The detailed knowledge of a pupils' profile of mathematical abilities was certainly absent when teachers were interviewed, yet the interviews did not provide an appropriate means of discovering reasons for this lack of adequate knowledge. There was, however, no criticism of the items on the checklist from the teachers, but there was a good deal of concern that the checklists were difficult to fill in, indeed there was some alarm from teachers that the constraints of time and syllabus did not allow the broad spectrum of abilities to be challenged. It was suspected that a reasonable O-level grade could be achieved by drill and exposure to appropriate questions, where 'tricks' could be learned that would lead to success. Because of the constraints under which the teaching proceeded this was a

73

plausable route to O-level, rather than a strategy that aimed for breadth and from which O-level would naturally emerge.

The most appropriate research method, therefore, was to work closely with the teacher during this second phase of checklist usage and, by means of classroom observation, investigate the problems at first hand.

The first exercise was to develop the checklist into a better instrument. A careful re-assessment of the items was made with the co-operation of both teachers from the schools and tutors from the Oxford Department of Educational Studies.

The aim of the development work was to ensure that the checklist items were appropriate for observation at third form level as well as valid for the identification of pupils able in mathematics. A careful check on the validity of these items was carried out. This work is described fully in Appendix 7.

An adequate amount of time was agreed with each of the teachers who would co-operate with the work, so that one of the research officers would be allowed to observe lessons over that period. Both teacher and observer would record clues to ability as they arose in the course of the classroom interaction with pupils. At an appropriate time teacher and observer would accumulate the information on a master chart and attempt an identification of the top 10% of the year group. After this exercise was complete a reassessment would be made and action taken to improve the method if it proved inadequate. This would be followed by a second phase of observation. Two terms were allowed for the initial planning, one for each phase, but it was found to be possible to complete the work in two half terms.

The question of reliability of observations is important in work of this kind so proper measures were taken to include reliability checks. The observer used a special form for recording details of the clues to ability during classwork. The relationship of the checklist items to a sample of clues was then discussed with an experienced mathematics teacher. This teacher also observed a sample of the same lessons as the research officer and records of observations were compared. Examples of the observations are included in Appendix 7, as are details of this reliability work, so no further discussion will be included here, other than to confirm that reliability of the observations was impressively high.

The question of reliability connected with scrutiny of workbooks was tackled in a different manner. We will see that very little could be confidently recorded across the spectrum of mathematical abilities when pupils books were 'marked' in the conventional way. The reliability of the evidence that was assumed from this aspect of the work was checked in co-operation with a mathematics tutor at the University of Oxford Department of Educational Studies. He was asked to list the challenges that were presented by a sample of exercises from the textbooks that pupils used and this was checked against the research officer's assessment. A high level of agreement was reached.

Added to these strict checks of reliability were the ongoing discussions of a more informal nature between teacher and research officer concerning the sorts of clues that the work with the pupils had revealed. There was complete agreement in this co-operative aspect of the work, though it must be stressed that records of clues picked up by teacher and observer were recorded separately. This was in order that an assessment could be made of the effectiveness of the teacher's observation, within the constraints caused by other aspects of the teaching situation, compared with the possibilities that could be exploited by an observer with just that task in mind.

A second reliability check was made after the 10% lists had been finally drawn up. Random samples of pupils were chosen for intensive testing in a one to one situation so that their profile of aptitudes could be adequately challenged, levels recorded and compared with the conclusions from the observational work. This is reported in Appendix 7 and the reader will see that there was no conflict between conclusions from this work and what was concluded from the observations.

We will now discuss results of the observational work. These will be given for the groups of pupils, one by one, and then we will draw together the overall conclusions.

In the checklist tables that follow the checklist categories are coded. It is appropriate to remind the reader of these codes at this point.

CHECKLIST OF MATHEMATICAL APTITUDES

UNDERSTANDING OF INFORMATION

1. Able to understand the formal structure of a problem, seeing the relationship between the items of information even when there is a lot to assimilate.

2. Able to see when there's redundant information in a problem even when there is a lot to assimilate.

3. Able to see when there is not enough information in a problem and to see what information is needed for its solution even when there is a lot to assimilate.

PROCESSING OF INFORMATION

4. Able to suggest plausible ideas that might lead to the solution of a new sort of problem.

5. Able to follow the logic in a long mathematical proof.

6. Able to construct mathematical proofs:
   a) similar to ones already demonstrated;
   b) requiring different logic from that already demonstrated but on a familiar topic;
   c) on problems related to a new or unfamiliar topic.

7. Quick to see when a method of solution that has been applied to problems in one topic will solve problems of a similar nature in another:
   a) with help;
   b) independently.

8. Quick to see when examples of a mathematical result have a wider and more general application: focuses readily on these generalisations rather than on the particular examples.

9. Able to follow a logical explanation or proof when valid logical jumps are made that miss out or assume some of the intermediate steps of logic.

10. Displays an ability to make valid logical jumps when using mathematical reasoning accompanied by the ability to provide the intermediate steps.

11. When a method fails in a process of reasoning is able to try an alternative approach incorporating what has already been validly reasoned into the new structure:
    a) in following someone else's logic;
    b) in constructing a system of logic independently.

12. Is able, when possible, to give alternative methods of solving a problem.

13. Is critical of inelegant solutions and strives in his/her own work for the most elegant.

14. Is quick to notice an incorrect step of logic or an invalid conclusion.

15. Is not content unless he/she has complete understanding of a mathematical concept, method or process of logic.

16. Adapts readily from a direct to a reverse train of thought.

MATHEMATICAL MEMORY

17. Has a good memory for a method of solution. Remembers a proof by its logic rather than by rote learning.

18. Has a good long-term memory for techniques developed for a particular topic. Requires little or no revision to regain previous levels of competance.

GENERAL

19. Shows little mental fatigue even after intense concentration on difficult problems.

20. Is not content to leave an interesting problem unsolved, but keeps returning to it at odd moments until success is achieved.

21. Can perform any arithmetical computation accurately and with ease.

22. When attempting to solve new sorts of problems displays a strong tendency to interpret and solve the problem in:
    a) a spatial/geometric way; or
    b) an analytic/logical way.

For the purposes of identification these were recoded according to the aptitudes that emerged from the work of V.A. Krutetskii (1976). The link between the codes is given in Appendix 7, and also on the relevant tables.

1. Obtaining mathematical information
   A. The ability for formalized perception of mathematical material, for grasping the formal structure of a problem.

2. Processing mathematical information
   A. The ability for logical thought in the sphere of quantitative and spatial relationships, number and letter symbols; the ability to think in mathematical symbols.

   B. The ability for rapid and broad generalization of mathematical objects, relations, and operations.

   C. The ability to curtail the process of mathematical reasoning and the system of corresponding operations; the ability to think in curtailed structures.

   D. Flexibility of mental processes in mathematical activity.

   E. Striving for clarity, simplicity, economy, and rationality of solutions.

   F. The ability for rapid and free reconstruction of the direction of a mental process, switching from a direct to a reverse train of thought (reversibility of the mental process in mathematical reasoning).

3. Retaining mathematical information
   A. Mathematical memory (generalized memory for mathematical relationships, type characteristics, schemes of arguments and proofs, methods of problem-solving, and principles of approach).

4. General synthetic component

A. Mathematical cast of mind.

Class 081

For the first half term the class observer played a passive role. By this we mean that he sat inconspicuously towards the back of the classroom with a seating plan and record sheet at hand so that every clue to ability that was available was recorded with the pupils name. Clues were classified as positive or negative according to whether the child showed a particular strength or weakness and were graded as High, Medium or Low. It was possible to categorise, in this way relative to class norms. For example we could establish where one pupils succeeded and another failed on similar work.

The teacher kept records of his/her own as well as he/she could. Lessons were usually in two parts, the first being class teaching where opportunities occured for pupils to show abilities in response to questions from the teacher or from their own questioning, and the second part being where the pupils worked on their own, attempting to solve mathematical problems based on the class discussion. During the second activity pupils would put up their hands to ask questions, usually when they were in difficulty. The teachers would give help as it was required. The observer continued the 'passive' role and supplemented what the teacher could do by answering some of these questions from individuals, recording any clues that were revealed. In this way we were able to record clues as they would naturally occur in the lessons.

Table 12.3 shows the levels that were confidently recorded by the observer during the first phase and Table 12.4 shows the records accumulated by the teacher over the same period (half a term) These tables can be compared with each other, and it is immediately apparent that an observer, with only one task in mind can record much more from the clues that come available than the teacher who is pre-occupied with all the other aspects of class activity.

TABLE 12.3

Checklist levels recorded by observer during 'passive' phase of observation (11 lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•			•	••		••													••		••					
08102						•					••											••					
08103				–	•	•			•		•						•	–		–							
08104				••			••			•	•							•		•					–		
08105							•				•											•					
08106	•			•									•														•
08107																					•	•					
08108	••			••	•	•	••		•	•	••		•			•	•		•	•	•				••		
08109						•	•				•									•							
08110						•																			•		
08111				•	•	•			•		–			•				•	•							•	–
08112						•					•									–		–			•		
08113	•			•						•	•			•													
08114						•					•									•							
08115						•					•			•				•				•					
08116						–				•										–	•	•			–		
08117	–			–			–				–							–			–	–	–				
08118						•				•	••												–				
08119				•	•						••							•		•							
08120	•			••	•		•			••	••	•	•	•				•	••	•	••	••	••				
08121				•		•	•				•							•				••					
08122	•			•	•	•														•							
08123	•			•	•		•		•	–																	
08124	•				•	•	•				•									•							
08125	•			•	•	•	•			•																	
	40	0	0	56	40	60	44	0	16	32	64	4	12	16	0	4	8	28	12	44	20	40	12	0	20	4	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.4

Checklist levels recorded by teacher during 'passive' phase (approximately half terms lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•			•					•	•	•											•					
08102																											
08103	•	•																•	•								
08104	•			•					•	•	•								•					•			
08105																											
08106																											
08107																											
08108		•				•	•		•	•								•	•					•			
08109																											
08110																											
08111				•										•	•												
08112				•																				•			
08113	•			•																							
08114																											
08115																											
08116																											
08117																											
08118																											
08119																								•			
08120				•							•		•	•	•									•			
08121	•	•		•					•	•	•		•							•				•			
08122																											
08123	•	•							•	•	•		•														
08124	•	•							•	•																	
08125																											
	28	20	0	28	0	4	4	0	24	24	20	0	12	8	8	0	0	8	12	4	0	4	0	24	0	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.5

Checklist levels recorded by observer after 'active' phase of observation (four further lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6 a	6 b	6 c	7 a	7 b	8	9	10	11 a	11 b	12	13	14	15	16	17	18	19	20	21	22 a	22 b
08101	•			•	••		••		•	•	••	•	•	•				•		••		••					
08102	•			•	•	•			•			•		•	–			•		•		••					
08103	•			–	•	•	–		•		•						•	–		–							
08104	•			••			••			•	•							•		•					–		
08105	•				•		•		•		•		•	•			•		•	•		•					
08106	•			•	•				•	–	•		•	•			•		•								•
08107	•			•	•				•	–			–						•		•	•					
08108	••			••	••	••	••	•	•	•	••		•			•	••		•	•	••		••		•		
08109	•			•		•	••			•	•	•	•					•		•			•				
08110	•					•	••			•										•		•			•		
08111	•			•	•	•	•		•		•		•	•			•	•	•							•	–
08112	–			•	•	•	•				•			•					–	–		–			•		
08113	•			•	•					•	•			•						–							
08114	•			•	•	•			•	•	••		•		•	•		•		•				•			
08115	•				•	•	•			•	•			•				•		•		••					
08116	•			–	•	•	–			•	•	–	–	•	•		–	–	–	–	•	•		–	–		
08117	–			–	–		–				–	–	–					–		–	–	•	–	–			
08118	–			–	–	•				–	••							–	–	–			–	–			
08119	•			•	•		•				••	•	•		•			•	•	•						•	
08120	•	•	•	••	••		•			••	••	•	•	•				•	••	•	••	••	••	•			
08121	•			••		•	••				••		•				••	••	•	•	•	••		•			
08122	•			•	•	•					–		–						–	••							
08123	•			•	•		•		•	–	–						•	–	•						•		
08124	•				•	••	••				••		•	•				••	•	•				•			
08125	•			•	•	•	•			•	–		–		–	–		–					–	–			
	100	4	4	88	84	60	72	4	40	64	88	28	64	44	20	12	32	68	56	80	76	44	24	32	24	8	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.6

Additional clues recorded by teacher during 'active' phase (three weeks)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101							•		•												•						
08103	•													•							•						
08104				•	•		•						•						•								
08108	•			•			•						•								•						
08112							•				•	•															
08113	•			•	•	•																					

TABLE 12.7

Checklist levels recorded from work book scrutiny, accumulated over 'active' and 'passive' phases.

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
08101	•					••	••			•										•		•					
08102	•					••	••			•										•							
08103	•					••	•			•																	
08104	•					••	•			•										•		•					
08105	•					••	•																				
08106	•					••	•			•										•		•					
08107	•					••	•			-										-							
08108	•					••	••													•		•					
08109	•					••	•			•			•							•							
08110	•					••	•			•										•							
08111	•					••	•			•										•							
08112	•					••	•			•										•							
08113	•					••	•			•										-							
08114	•					••	•			•										•					•		
08115	•					••	•			•										•		•					
08116	•					••	•			•																	
08117	•					••	-			•										•							
08118	•					••																					
08119	•					••	••			•										•							
08120	••					••	••						•							•		•					
08121	••					••	••			•										•		•					
08122	•					••	•			•										•							
08123	•					••	•			•										•							
08124	•					••	••			•										•		•					
08125	•					••	•																				
	100					[illegible]	96			[illegible]																	

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.8

Recategorisation of information from Table 12.4 to Table 12.7
for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2, 3)	2A (4,5, 6,21)	2B (7,8)	2C (9,10)	2D (11,12, 13,19)	2E (14, 19, 20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
[illegible]	•	••	••	•	•	•	••	••		Top 5%
[illegible]	•	••	•	•	•	•	•	••		10%
[illegible]	•	–	•		•	–	–	•		
[illegible]	•	••	•			•	•			
[illegible]	•	•	•	•	•	•	•	•		10%
[illegible]	•	•	•	•	•	•			[illegible]	10%
[illegible]	•	•	–	–		•	–	•		
[illegible]	••	••	••	•	••	•	••	••		5%
[illegible]	•	••	•	•		•	•			10%
[illegible]	•	••	•				•	•		
[illegible]	•	•	•	•	•	•	•		[illegible]	10%
[illegible]	–	•	•	•		–	–			
[illegible]	•	•	•		•		–			
[illegible]	•	•	••	•	•	•	•			10%
[illegible]	•	•	•		•	•	•	••		10%
[illegible]	•	–	•	–	–	–	–	•		
[illegible]	–	–	•	–	–	–	–	–		
[illegible]	–	–	•			–	–			
[illegible]	•	••	••	•	•	•	•		[illegible]	10%
[illegible]	••	••	••	•		••	•	••		5%
[illegible]	••	••	••	•	••	••	•	••		5%
[illegible]	•	•	–	–	•	–	•			
[illegible]	•	•	–		•	–	•			
[illegible]	•	••	••	•	•	••	•	•		10%
[illegible]	•	•	–	–	–	–				

Furthermore, the grading of observations that was possible for the observer was not possible for the teacher. The coding for tables based on the observer's records is ** = Grade H(high), * = Grade M(medium), - = Grade L (low). The single code for tables based on teachers records is: * = 'showed strength on this category'.

It is important to note that from the sample of classes observed, Class 081 produced the highest rate of clues. The teacher had a teaching style that involved a large proportion of the class in questions and answers as a topic was developed. Table 12.3 can be compared with later tables of a similar kind based on other classes. It is apparent, however, that it was not possible for the teacher himself to maintain adequate records.

The profiles gradually emerged for a number of pupils and it could be conjectured that given more time Table 12.6 would be closer to completion. There were some checklist items that were not observable at all during this phase and there were a considerable number of pupils on whom decisions could not be reached on a large number of items.

Table 12.3 provided a sound basis for the second phase of activity. This is where the observer played a more 'active' role. At appropriate times, when the pupils were working independently, the observer would spend a few minutes with pupils on whom certain categories of information has not been recorded. He would chat to them about the work and pose questions that they had to tackle on the spot, related to the levels of attainment on the current work. In this way further clues could be recorded. The observer was merely simulating what a teacher could do during a lesson. This active questioning by the teacher and positive feedback must be compared with the answering of a pupil's questions when they get into difficulty - a process resulting in negative feedback.

It was possible for the teacher to pursue this course of action to some extent at the same time, but for the purposes of the experiment the observer did most of the work.

As a result of the 'active' phase, Table 12.3 was updated to Table 12.5. Extra clues that the teacher recorded during this time appear in Table 12.6. Of course, at this stage, there were many more gaps in the teacher's own records than in the observers.

Table 12.5 evolved into a tool from which reliable identification could be made. Thus this teacher had a style which allowed adequate clues to be gathered from the classroom activity alone, even though he himself could not keep adequate records.

There were still some checklist items that could not be recorded, however. Items 2 and 3 - the ability to handle problems with redundant or insufficient information, item 6C - ability to solve problems of an unusual nature, item 12 - the ability to look for alternative methods and item 22 - the cast of mind, were all very scantily observed. This implies that, even in the 'active' phase, the type of the challenges set to the pupils did not allow conclusions to be reached on these items. In particiular, the lack of challenge coded as 6C is of concern when one considers the adequacy of high level challenges for children of high ability.

Table 12.7 records the checklists levels that were discovered from scrutiny of workbooks. The very narrow range of possibilities is quite apparent. Much of the work that pupils present in their workbooks is of a routine nature, either solving problems of a kind that the teacher has demonstrated in class (item 6a) or one on the same topic with just minor refinements of technique (item 6b). Not all of the pupils progressed to the 'hardest' questions, which could usually be categorised as 6b H; they could get by on the drill questions. When harder questions are attempted, it is not easy to associate the relevant checklist items with the pupils' work. First it would take a good deal of time to put the clues together for each child, and second, when the pupil is not present to discuss the work, it is difficult to deduce the route that led to what was finally written in the workbook. How long did he spend on it? Did he have help? Did he try different approaches?

It is important to consider, however, that most of the conclusions that teachers draw are based on records of attainment in workbooks. In this particular class we can tell from Tables 12.4 and 12.6 that it was not possible to draw valid conclusions from records of clues in the classroom. In contrast, the clues from the classroom, that the observer was able to record were both more reliable and more adequate than records from workbooks.

Table 12.8 shows the recategorisation into items based on the original source (the work of V.A. Krutetskii). We recall that various checklist items referred to similar Krutetskii categories. We should note that a broad spectrum of items were observable from pupils in this class, who were from the 13-14 year age group. Nevertheless some of the Krutetskii category levels could only be determined from a limited range of checklist items. For example, Krutetskii's Category 1, 'Gathering of Information', could only be determined from clues related to one of the three possible items on the checklist (1,2,3) to which it was related. We can conjecture that it was the kind of mathematical provision which dictated this limitation, so that though relative abilities in this group of children showed up - the identification of top 10% and 5% is given in Table 12.8 - the differences would be more marked if higher levels of challenge had been present.

School 03

The work in school 03 adds to our insights in a number of ways. First, we are dealing with a different kind of teaching environment. This is important in view of our discovery that differences of accuracy and bias of judgement do exist from teacher to teacher within a subject. Second, in school 03, we had the co-operation of teachers of both the top set and second set in mathematics. This made it possible to try out the checklist and assess its effectiveness as an aid to classification in two groups of pupils for whom there were different levels of expectation. This allowed us to assess the useability of the instrument in these two settings and also to consider the identification of the top 10% over the combined group. It may have been that some pupils of high ability were in the second ability set.

The conclusions are connected with passive observation of a sample of lessons over half a term on the part of the observer and on the accumulation of clues on the part of the teacher, over the same period, that emerged from the normal classroom routine.

We will report the results for the two classes separately.

Class 031 (The higher ability set)

Tables 12.9 to 12.12 contain the checklist levels that were recorded over this period. The style of teaching differed from the previous one reported (class 081). The class teaching and discussion was much more formal so that interaction between pupils and teacher were more limited. The class was certainly questioned during the discussion of the topic, but the bias was towards formal instruction rather than involved discussion when the class were actively encouraged to contribute to the development of the

88

work. The rate of observation of clues during this phase of activity was thus lower for two reasons. First the class involvement was not so high and second because, this formal part of the lesson was shorter, being used as a device for moving the pupils onto exercises that they attempted individually.

During the second phase of the lesson the teacher answered individual problems that pupils had, but also found time to do some marking of work from which discussion could emerge.

The observer was able to do the same as the teacher during this second phase and it was from here that the majority of clues emerged. Table 12.9 can be compared with Table 12.3. The majority of clues for Table 12.9 came from the second phase of the lesson. The majority for Table 12.3 came from the first phase of the lesson.

The teaching styles produce similar amounts of clues, though the kinds in Table 12.3 associated with class discussion are more variable in nature than from Table 12.9 associated with individual discussion. Table 12.9 gives complete information for the class in certain categories. These are rather like the categories that can be observed when workbooks are scrutinised. These categories do, however, link with the kinds of questions that pupils are asked to attempt, limited in this class to drill and practice based on the methods demonstrated in the first part of a lesson.

TABLE 12.9

Checklist levels recorded by observer from class observation ('Passive' ten lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03101	•			•		•	•		•	•	•											••			•		
03102	•			•	•	•	••		•	•	•									•					•		
03103	–					•	–		•											•					•		
03104	•					•	–		•	•	•									•					••		
03105	•					••	•		•	•	•			•						•					•		
03106	–					•	–		•											•					•		
03107	••			•	••	••	••		••	•	••							•	•	•		••			••		
03108	•					•	–		•											•							
03109	•					•	••		•	•	••									•					•		
03110	•			•		•	•		•	•								•		•					•		
03111	••					•	•		••	•										••					•		
03112	•			•		•	–		•	•	–							•		•		•			•		
03113	•			•		•	–		•	•										•		•			•		
03114	•			•		••	•		•	•	•									•					•		
03115	••	•		•	••	••	••		••	•	••		•				•		••	••		•			••		
03116	–					–	–		•	•	•					–			–	–					•		
03117	•			•		–	–		•									•		•		•			•		
03118	•			•		•	•			•	••									•							
03119	•			•	•	•	•		•	•	••		•							•					•		
03120	•			•		••	•		•	•	••									•					•		
03121	•					•	•		•											•		•			•		
03122	•			•		••	•		•	•	••									••					•		
03123	•					•	•		–											–							
03124	•					•	•		•		•									•					•		
03125	•					•	•		•											•					•		
	100	4	0	52	16	100	100	0	96	68	60	0	8	0	0	4	4	16	12	100	0	28	0	0	88	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.10

Checklist levels recorded by teacher (Approximately half term)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03101									●										●		●						
03102		●			●	●	●			●	●	●		●	●				●		●		●		●		●
03103																											
03104						●			●		●		●						●		●		●	●	●		
03105																					●		●	●	●		
03106																			●						●		
03107		●			●	●	●		●	●	●	●	●						●	●	●		●	●	●		●
03108																											
03109									●												●				●		
03110				●									●								●				●		
03111		●								●									●		●			●	●		
03112																									●		
03113				●									●												●		
03114																			●						●		
03115		●			●	●	●			●	●	●	●						●	●	●		●		●		
03116											●								●						●		
03117																											
03118									●		●								●						●		
03119		●							●												●				●		
03120																			●						●		
03121									●																●		
03122		●			●					●	●								●		●				●		
03123																											
03124						●	●		●	●						●			●		●		●		●		
03125																									●		
	0	24	0	8	16	20	20	0	32	24	28	12	20	4	4	4	0	0	52	8	48	0	24	16	80	0	8

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.11

Checklist levels recorded from workbook scrutiny checklist codes

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03101						•				•																	
03102	••			•		••	••				•									•							
03103						•																					
03104						••	•													•							
03105						••	•													•							
03106						•	–													–							
03107	••					••	••			•										•							
03108						•	–																				
03109	•					••	•			•										•							
03110						••	•																				
03111	•					••	••													•							
03112	•					••																					
03113						•	•																				
03114						••	•																				
03115	••					••	••			•										•							
03116						•														–							
03117						•																					
03118	•					••	•			•										•							
03119	•					••	•						•														
03120	•					••	•													•							
03121	•					••	•													•							
03122	•					••	••													•							
03123						•	–													–							
03124						••	••													•							
03125	•					••	•													•							
	48	0	0	4	0	100	80	0	0	20	4	0	4	0	0	0	0	0	0	64	0	0	0	0	0	0	0

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

TABLE 12.12

Recategorisation of information from Table 12.9 to Table 12.11 for the purposes of identification of top 10% and top 5%

Pupil Code	1A (1,2,3)	2A (4,5,6,21)	2B (7,8)	2C (9,10)	2D (11,12,13,19)	2E (14,19,20)	2F (16)	3A (17,18)	4 (22)	IDENTIFICATION LEVEL
03101	•	•	•			•		••		Top 10%
03102	•	••	•	•	•	•	•	•	ANAL/LOG	Top 5%
03103	–	–	•			•	•			
03104	•	–	•	•		•	•	•		
03105	•	•	•		•		•	•		Top 10%
03106	–	–	•			•	–			
03107	••	••	••	•	•	•	•	••	ANAL/LOG	Top 5%
03108	•	–	•				•			
03109	•	•	•			•	•			
03110	•	•	•	•		•	•			
03111	••	••	••			•	••			Top 5%
03112	•	–	–			•	•	•		
03113	•	–	•	•			•	•		
03114	•	•	•				•			
03115	••	••	••	•	•	••	••	•		Top 5%
03116	–	–	•		–		–	–		
03117	•	–	•			•	•	•		
03118	•	•	••				•			Top 10%
03119	•	•	••	•			•	•		Top 10%
03120	•	•	••			•	•			Top 10%
03121	•	•	•				•	•		
03122	•	••	••			•	••	•		Top 5%
03123	•	–	–				–			
03124	•	••	•		•	•	•	•		Top 10%
03125	•	•	•				•			

Categories 6a and 6b are associated with these routines. Category 1 is associated with the ability to understand information, but the abundance of single *'s in this column indicates that high level challenges were not in general presented. Item 21 concerns aptitude in arithmetic - again not challenged at a high level. One might suppose that the clues could be much stronger and variable either if the challenges were more adequate and variable in the work set to the class as a whole or in the discussion between individual pupil and teacher.

It is interesting that the amount of information that the teacher himself recorded was much fuller than in class 081 and this is a consequence of the time devoted to work with individuals. There is a little more variety in Table 12.10 than Table 12.9, but there are signs in both tables of more complete knowledge of some pupils than others.

The same conclusions about inadequacy of the clues that can be used from workbook scrutiny is apparent from Table 12.11.

If one were to attempt a classification of pupils into top 10% and next 10% (in this case top and bottom halves of the class) the lack of the full range of clues, particularly from the teacher's own records would make the exercise tentative, yet when the three tables are combined and reclassified into Table 12.12, we see that from this 'passive' activity a preliminary classification was possible. If we had then pursued an 'active' phase where we worked closely with pupils on whom knowledge was rather scant, we could improve the table of results and carry out a more confident classification.

It is important to repeat that the range of clues that combined to give the categorisations of Table 12.12 was limited by the level of provision. There must be a source for concern if the sort of work presented to pupils does not challenge the full range of aptitudes, both because of mathematical growth of the children and of their identification. We must remember too that the tentative identification that was possible from Table 12.12 relied heavily on the categories discovered by the observer. Thus there were enough clues possible for an identification but the teacher had been too preoccupied to record a sufficient number.

Class 032 (Second ability set)

The style of teaching of this group was much the same as for 031. The marked difference was in the level of challenge set to the pupils. The teacher was at pains to make sure that all pupils could gain a working knowledge of the topics covered[1]. This inevitably led to an algorithmic approach and to routine exercises. This was understandable in that many of the class achieved success on nothing but routines.

The formal part of the lesson did allow for some variety of clues to be picked up because the teacher looked to the class for ideas and involved a large number of pupils.

The classification of clues was relative to the levels of challenge set within this group, so that norms were different from that for class 031. For this reason it was difficult to establish cross group norms, though it did appear that the best of class 032 would not be a candidate for the top 10%.

The exercise was to try to divide the class into two equal groups on mathematical ability. In this way we could establish whether the checklist was an effective instrument in this lower ability environment.

The clues that could be observed were inevitably of the kind associated with the routine work that the children were set and these routines made it impossible to get very much from workbooks. The dominant item here was 6a, associated with solving problems by means of an algorithm. There were very few clues to any other aspect of mathematical ability, including the ability to understand information. By way of contrast, Item 1 was usually observable from the workbooks of higher sets.

Again the teacher recorded an inadequate amount for proper identification, but added to the observer's records and when all clues were reformulated into Table 12.16 we see that a spectrum of abilities could be built up for a number of pupils.

TABLE 12.13

Checklist levels recorded by observer from classwork
(Passive role, Ten lessons)

CHECKLIST CODES

Pupil Codes	1	2	3	4	5	6a	6b	6c	7a	7b	8	9	10	11a	11b	12	13	14	15	16	17	18	19	20	21	22a	22b
03201	*			* *	* *	* *	* *			*	* *					*				* *					*		
03202	*			*	* *	*					* *									* *					*		
03203	*					–				–	–														*		
03204	*					*	–		*	–									–					–			
03205	*					*	*																		*		
03206	*			*	–	*	–				–									–					*		
03207	*				*	* *	*			*	* *									*					*		
03208						*	*		*	–										*					*		
03209	*			*		–														–					–		
03210	*			* *	* *	* *	*				*								*	*					*		
03211	*			*	–	–	–				–									–					–		
03212						*	*																		*		
03213	* *			* *	* *	* *	* *						* *			* *	* *			* *					* *		
03214	*			*	–	–	–				–									–					–		
03215	–			*	*	–																			*		
03216	*				*	*	*				* *									* *					*		
03217		*			–	–					–									–					*		
03218				*	*																						
03219	*			* *	–	–	–				–									–					–		
03220	*				*	* *	* *				* *									*					*		
03221				*		–					–		*			*	*					–					
03222	*			*	–	*	*				–														*		
03223				*		*	*	–												–							
03224	–				–	–	–				–												–				
03225	*			*	*	*	*			–								*		*					–		*
03226	–				–	–	–				–						–			–					–		
03227	*				*	*	*													*					*		
	78	4	0	56	70	96	70	4	7	19	59	0	7	0	0	7	11	4	7	67	0	4	4	4	81	0	4

PERCENTAGE OF PUPILS WHO SHOWED ADEQUATE INFORMATION

	1 Always	2 Often	3 Sometimes	4 Rarely
6. oral response alert oral response to questions in French (may be superior to written response).				
7. sound/symbol correspondence when reading aloud pronunciation not perverted by written form of language.				
8. self-confidence i) does not appear concerned about looking foolish when speaking French.				
ii) enjoys trying out French sounds.				
ATTITUDE (while high scores on the following items in contrast to low scores on previous items do not necessarily indicate high ability, low scores in contrast to previously high scores may indicate under-achievement).				
9. enjoyment shows enthusiasm for the study of French.				
10. relevance of French acknowledges relevance of French to own future needs.				
11. interest in French/foreign culture shows interest in learning about French way of life/the life of foreign countries generally.				

PHYSICS CHECKLIST (YEAR 1)

NAME OF PUPIL ....

The boxes 1-4 on the checklist are to be ticked according to the following key:-

1) This item describes this pupil's behaviour on all or almost all occasions.
2) This pupil frequently behaves in the way described.
3) This pupil sometimes behaves in the way described but has difficulty in sustaining this level of work.
4) This pupil rarely or never displays the behaviour described.

The one word headings to the column could be misleading on their own and should be regarded just as memory aids.

For some items indications are given, in brackets, of areas of the syllabus where such behaviour may be observed.

	1 Always	2 Often	3 Sometimes	4 Rarely
A. HANDLING APPARATUS 1. Is able to use new apparatus with little guidance but is aware of his limitations and will ask for help if necessary (use of microscopes or ticker timers)				
2. Treats apparatus with appropriate respect				
B. MAKING OBSERVATIONS 1. Makes observations without any form of guidance and can select the relevant from the non-relevant (ripple tanks)				
2. Does experimental work with care and has a feeling for the magnitude and consequences of any remaining error				
C. LOOKING FOR PATTERNS 1. Attempts to incorporate new observations with previous ones, into a pattern (circuit board experiments)				
D. COMMUNICATING 1. Chooses appropriate methods of recording results (diagrams, tables, graphs, descriptive writing)				
2. Is accurate in the technical use of common words (e.g. can distinguish between common and scientific uses of words like force, energy)				

	1 Always	2 Often	3 Sometimes	4 Rarely
E. EXPLAINING AND PREDICTING 1. Is persistent in seeking to generate explanations - mere descriptions will not do. May provide several explanations and rank them in order of plausibility				
2. Seeks to make predictions but remains receptive to unexpected results. ("Monkey and Hunter" expt.)				
F. TESTING IDEAS 1. Seeks to test ideas, (predictions, explanations) experimentally and will reformulate the problem if necessary to make this possible.				
2. Can make decisions about which variables to control and which to investigate (investigation of springs or levers)				
G. MATHEMATICS 1. Is accurate in manipulating numbers and symbols				
2. Does rough calculations with an accuracy appropriate to his/her purpose				
H. CONCEPTS 1. Learns new concepts with ease (current, acceleration)				
2. Can apply appropriate concepts to new problems				

	YES	NO
I. GENERAL 1. Can see the direction of an argument, anticipate its outcomes and grasp its implications		
2. Has "scientific" interests		
3. Relates school science to the world outside the lab. - is keen to know how things work		
4. Can detect flaws in the logic of an argument		

This checklist is based on the Assessment of Performance Unit "Science process list", the Cleveland checklist for the assessment of ability in science and the H.M.I. report on science in "Gifted Children in Middle and Comprehensive School". Close discussions with a large number of teachers have been a feature of all of these pieces of work.

NAME OF PUPIL ______________________________

CHECKLIST OF MATHEMATICAL ABILITIES CHARACTERISING THE MORE ABLE (YEAR 1)

A. UNDERSTANDING OF MATHEMATICAL INFORMATION

A pupil who has high ability in a branch of mathematics must be able to see the relationships between the items of information in a problem. He sees the situation as a whole rather than its separate parts. He is able to see what information is essential and what is superfluous in a problem, and what sorts of conclusions can be drawn from the information.

	Rarely if ever	sometimes	On most occasions
1) Is not put off by the amount of information in a problem.			
2) Makes a quick start to solve problems even when there is a lot of information to absorb.			
3) Is able to work completely independently on project work that requires the gathering and sorting of data.			

B. PROCESSING OF MATHEMATICAL INFORMATION

Speed in solving problems is not a good diagnostic trait. The mathematician is characterised more by his speed in understanding concepts and logic. He may be slow at solving problems for such reasons as striving for elegant or unusual solutions.

4) Is able to work accurately in all operations with numbers. (It is not necessary for him to be able to do large calculations in his head).			

Ability to generalise is an important quality.

5) Is able to work and think fluently when symbols replace numbers.			
6) Is able to transfer rapidly from concrete cases to generalisations.			
7) Is able to adapt a method of solution to unusual situations (e.g. tackles the 'miscellaneous' exercises with confidence).			

8) Is able to make logical jumps in a solution as if by inspiration. (Some difficult steps may seem 'obvious' but hard to explain)			
9) Is able to follow the logic in long proofs with ease.			
10) When one method fails, is quick to look for another.			

11) Works fluently with geometric transformations			

Rarely if ever | sometimes | On most occasions

12) Copes easily with 3 Dimensional Problems.

13) Displays sudden inspiration in finding a method to solving difficult or unusual problems.

14) Is not content with a solution unless the logic is sound.

15) Is not content with an inelegant solution. Strives for the neatest and clearest method.

C. INFORMATION RETENTION

A Mathematician may have a bad memory for specific facts or data. He will, however, have a good memory for generalisations and patterns of logic. We may deliberately only commit generalisations to mind.

16) Has a good memory for method of solution.

17) Is quick to formulate a general strategy for solving a problem, drawing on previous experience.

D. GENERAL

18) Has a mathematical cast of mind. Tends to focus on the mathematical content of a situation. (e.g. Adds up car numbers on a journey, estimates the size of a football ground etc.)

19) Is not content to leave a problem unsolved. Is 'haunted' by it, coming back to it at odd moments until a solution is found.

20) Shows no fatigue in maths lessons even when engaged on hard concepts or difficult problems.

21) Shows great satisfaction in solving mathematical problems. (This may not have strong external expression or may be of quite short duration between problems).

22) Has a critical approach to his own achievement and potential in mathematics.

E. ATTITUDES

(Do not use answers in this section to help judge the more able mathematician. These characteristics can be misleading. High scores here and low scores in the main sections may shed light on wrong identification. Low scores here and high scores above, may support an underachieving but naturally able mathematician.)

1) Extremely perseverent in mathematical work.

2) Always completes the maths work that is set.

3) Works, always, with extreme care.

4) A very neat worker.

5) A very patient worker.

ENGLISH CHECKLIST (YEAR 2)

WRITING 1. Technical control i) writes complex sentences/ shows syntactical sophistication.											
ii) manipulates punctuation to ensure intended meaning.											
2. Reader Awareness Provides adequate context/background information in own writing: conscious of needs of potential readers.											
3. Variety of Register Able to adopt variety of registers/ styles as suitable.											
4. Humour/Irony Has sense of humour able to make connections revealing ironies/absurdity of a situation etc.											
5. Vocabulary Uses extensive resources of words: willing to experiment.											
6. Originality Strives for fresh, inventive plots/ situations in fictions.											
7. Sympathetic Imagination Able to imagine and write from wide variety of viewpoints.											

ORAL

8. High Oral Ability

(This refers to a pupil's use of speech in the classroom to discover, investigate and express ideas out loud - whether in class discussion, group talk or one-to-one conversation - and the extent to which the pupil listens and responds to what other people say.)

i) Talking											
ii) Listening											
READING 9. Comprehension i) Factual. Able to extract, select, synthesise facts from a passage/passages.											
ii) Meaning. Sensitive to subtle or implied meanings.											
10. Personal Reading i) Enthusiasm: is a keen and voracious reader.											
ii) Range: is willing to tackle a wide range of reading material (including unfamiliar or demanding).											
11. Reading Out Loud Can read out loud spontaneously with appropriate sensitivity and intonation.											
12. Critical Judgement Advances beyond purely subjective judgements: elaborates on own opinions by reference to text etc.											

FRENCH CHECKLIST (YEAR 2)

1. **Attitude** i) Shows enthusiasm for the study of French/perceives the subject as relevant to own future needs.												
ii) Shows active interest in learning about French way of life/the life of foreign countries in general.												
2. **Aural/oral skills** i) Able to discriminate between foreign sounds.												
ii) Able to articulate foreign sounds.												
3. **Oral response** Alert oral response to questions in French (may be superior to written response).												
4. **Control over sound/symbol correspondence** Pronunciation not perverted by written form of language.												
5. **Self-confidence** Not embarrassed when speaking French/enjoys trying out French sounds.												
6. **Memory** Shows clear evidence of good memory.												
7. **Mastery of English** i) Knowledgeable of essential grammar e.g. parts of speech.												
ii) Alert to nuances in own language.												
8. **Flexibility** i) Can adapt to an entirely new set of rules, and can think within them.												
ii) Does not try to impose English style or syntax on French sentences when writing compositions.												
9. **Ability to put the language together** Independently makes new connections out of isolated units of knowledge - makes creative use of French structures.												

SCHOOL 09 PHYSICS CHECKLIST (YEAR 2)

Description of Items

A. Handling Apparatus

1. Is able to use new apparatus in an appropriate way with little guidance.

2. Is able to explore the possibilities presented by a new piece of apparatus but is alert to its limitations.

(In A.1 we are thinking of the sort of person who can use a ticker-timer for a standard experiment with little guidance on such matters as how to connect it up and where to put the tape. In A.2 we were intending to focus on a rather different skill - for example, the ability to recognise that there is a range of speed for which the timer would be appropriate but that, driven from a $50H_z$ source, it could not be used to measure very slow speeds. "Exploring the possibilities" may be done by means of small scale experiments (e.g. trying to pull a tape very slowly to see what happens) or by discussion with other pupils or the teacher).

B. Making Observations

1. Makes observations without any form of guidance and can select the relevant from the non relevant.

2. Does experimental work with care and has a feeling for the magnitude and consequences of any remaining error.

(B.1 might be noticed when pupils are working with ripple tanks. Some find it difficult to see the incident waves and the reflected, refracted or diffracted waves amongst the inevitable background of spurious reflections, even when the teacher is pointing them out. Others can see these main features without any help. This ability to make observations and to be selective is the clue to high ability that we were intending to focus on. Just being able to make observations without help doesn't seem to be a strong enough clue on its own. In B.2 we had in mind the sort of pupil who will take care to friction compensate a trolley board but who is able to appreciate that there may still be some friction (especially if trolleys are stacked up) which might explain otherwise unexpected results.)

C. Looking for Patterns

1. Attempts to incorporate new observations with previous ones into a pattern.

(Circuit board work or ray streak optics experiments provide opportunities for pupils to display this ability. We were thinking of the sort of person who can relate the effect of placing a resistor in a circuit to the effect of placing a second bulb in the circuit and perhaps suggest that bulbs have the same property that resistors have.)

D. Communicating

1. Chooses appropriate methods of recording experiments and results (diagrams, tables, graphs, descriptive writing, etc.).
2. Is accurate in the technical use of common words (e.g. can distinguish between the common and scientific use of words like force, energy, power).

E. Explaining and Predicting

1. Is persistent in seeking to generate explanations - mere descriptions will not do.
2. Seeks to make predictions but remains receptive to unexpected results.

(E1 may be recognised through the way that a pupil talks or writes about an experiment. It may also be indicated by the quality of the questions that he or she asks. Pupils seeking explanations may ask more "why?" questions - others who do not see the importance of explanation (or even the difference between description and explanation), may ask more "what happened?" questions. A pupil with a highly developed sense of the importance of explanation might generate a number of explanations and rank them in order of plausibility (this may not be written up but might be noticed in a pupil's talk about an experiment).

In E2 we were intending to focus on the pupil who does not overlook the result that does not fit in with his/her expectations but who does bring it to light and ask about it or comment on it.)

F. Testing Ideas

1. Seeks to test ideas (predictions, explanations) experimentally and will reformulate the problem if necessary to make this possible.
2. Can make decisions about which variables to control and which to investigate (e.g. in explanatory work with springs or levers).

G. Mathematics

1. Is accurate in manipulating numbers and symbols.
2. Does rough calculations appropriate to his/her purpose.

H. Concepts

1. Learns new concepts with ease (current, acceleration).
2. Can apply appropriate concepts to new problems.

I. General

1. Is not perturbed when attempts to solve a written problem fail - will try an alternative approach.

2. Has "scientific" interests.

3. Relates school science to the world outside the lab - is keen to know how things work.

4. Can detect flaws in the logic of an argument.

PUPILS' NAMES

A1. Appropriate use of apparatus - without help.														
A2. Alert to possibilities limitations of apparatus.														
B1. Selective in observing.														
B2. Careful experiments, aware of consequences of error.														
C1. Fits new observations into patterns.														
D1. Appropriate methods of recording.														
D2. Accurate use of technical terms.														
E1. Explains not just describes.														
E2. Predicts results but can accommodate surprises.														
F1. Looks for experimental test of ideas.														
F2. Can choose variables to control.														
G1. Accurate with numbers and symbols.														
G2. Does rough calculations where appropriate.														
H1. Learns concepts easily.														
H2. Can apply concepts to new problems.														
I1. Will try alternative approaches to problems.														
I2. Has scientific interests.														
I3. Relates school science to outside world.														
I4. Detects logical flaws.														

17 December 1981

Ref: OERG/MA/CD/DA

I enclose, as promised, the new version of the maths checklist and a master sheet. I should be grateful for any comments you may have before we put the method of identification to the test next term. At any rate you should feel free to use it in the best way to suit your own situation. The items will be relevant to the pupils' work at different times and some will be more observable than others, but over the term we should be able to decide which aptitudes we can reliably monitor by our observations of classwork, homework and any out of class activities.

I think that the following scheme will be workable.

1. Carry the single sheet checklist with you to all lessons and have it handy when marking homework. I think that the items will soon be remembered so that constant reference will not be necessary.

2. Wherever a cue to ability is shown make a brief note of who showed his/her potential and in what context alongside the item on the sheet. If it is possible you should try to assess the relative position within the class on the observed category. This may be difficult but would give a more meaningful picture towards the end of term. The position could be as coarse as 'top half of class' or as precise as 'within top three of class'. These estimates could well be revised in the light of later information. If this is too hard to assess it would be appropriate to put a tick as indicating 'shows high ability on this aptitude'. It would be useful, too, if we could indicate either the maths topic that revealed the aptitude or to indicate whether it was revealed in the content of numerical, spatial (geometric) or algebraic work.

3. If there are clear indications of a lack of these aptitudes a similar record should be made.

4. From time to time the master chart should be updated.

5. At the end of term the master chart should be consulted and an attempt be made to list which pupils are in the top 10% based on the combinations of abilities that they have displayed over the term. If the evidence is adequate further subdivisions should be attempted e.g. top 5% or top 1%. These percentages refer to percentages of the whole year group, so that if the top set is one fifth of the year group we would look for the top half of this set if we wanted to list the top 10%. If a rank ordering is possible this should be attempted.

Some of the checklist items require some amplification:

Item 1. This attempts to summarise the ability to understand how the information in a problem is linked together in such a way that deductions can be made from it. One might see this in pupils not being put off by the amount of information that a problem contains and making a quick start to solve problems even with a lot of information. Items 2 and 3 are further clues to this ability.

Item 8. For example we may detect an ability to move quickly from examples in arithmetic to their algebraic generalisations.

Item 16. For example, quick to see the converse of a theorem, or having been taught a result like $(a + b)^2 = a^2 + 2ab + b^2$, readily adapting it to factorization such as $4x^2 + 4x + 1 = (2x + 1)^2$. This can be detected by the ability to solve a 'reverse' problem, having just solved a 'direct' one. A simple example could be:

DIRECT How many days does it take to earn x pounds at y pounds per day?

REVERSE How much does a worker earn in a day if he earns b pounds per day.

Item 22 is present to monitor the particular bias of the pupils' ability rather than as an item for the purposes of identification. It is possible that the pupil will have a strong leaning to one or other of spatial or analytic (algebraic) 'casts of mind'.

Please get in touch if anything else is hard to interpret: we could get together and talk in more detail. I will get an adequate supply of these checklists to you for next term and will modify them by then if there are a number of new ideas that need to be incorporated.

Thanks again for your help in what I hope we will all find a worthwhile exercise.

Yours sincerely,

Cliff Denton

FORM ____________

SCHOOL

	UNDERSTANDING INFORMATION			INFORMATION PROCESSING																			MEMORY		GENERAL					
NAME	1. Understanding problems with a lot of information.	2. Seeing when information is redundant.	3. Seeing when there is not enough information.	4. Suggests ideas for new sort of problem.	5. Ability to follow a long proof.	6. Able to construct proofs:	a. similar to ones shown;	b. requiring new logic but on familiar work;	c. on new sorts of problems.	7. Able to apply old methods to new topics:	a. with help;	b. independently.	8. Able to transfer to generalisations quickly.	9. Following proofs with logical jumps.	10. Making logical jumps with understanding.	11. Trying alternative approaches aid-proof:	a. following a proof;	b. constructing a proof.	12. Finding several methods.	13. Striving for elegance.	14. Detecting invalid logic.	15. Strives for understanding.	16. Direct to reverse train of thought.	17. Memory for method or proof.	18. Long term memory of work covered.	19. Little fatigue.	20. Not content with unsolved problem.	21. Arithmetical skills.	22. Cast of mind:	a. spatial/geometric	b. analytic/logical

APPENDIX 6

Concerning the sample from which the teacher nominated the top 10%

In the bulk of our work we have based our conclusions on lists drawn up by teachers after due consideration of a sample of the school's third year group. On the whole this sample will not include every child in the year group. Pupils in low ability sets will not have been considered. It could be argued that by the third year some pupils of high ability will already have been undervalued and put into these lower ability sets and gone on to be unrecognised as they do no more than perform to expectations.

We discussed this to some extent in Chapter Four, pointing out that many of our conclusions were concerned with the pupils whom the school 'system' had identified, the teacher's list just being the final step in that 'system'. However, the evaluation of pupils that is done from time to time and upon which final setting depends, relies on teachers' judgement, either as a body or as individuals.

There is no doubt that some pupils would be missing from lists because of the above problem but are confident that the number would be relatively small, so that the main thrust of our conclusions must still be directed at those teachers who compiled the lists for us.

It has been possible to study one further aspect of this because of the fact that all of the pupils from whom the top 10% lists were chosen after the first year's checklist usage were named by the teachers. We show, in Table 1, the main variables that characterised the selection of the pupils. The method of analysis was similar to that reported in Chapter Four and relied on the technique of Partial Correlation.

Similar conclusions apply to the characteristics that are shown in this table to those found in our other analyses, reported fully in Chapter Four. The characteristics, which are in some way connected with teacher bias, do indicate the same level of uncertainty, that we discussed in Chapter Four because the characteristics are not usually the most appropriate for each subject in most schools. There is also much variation in the sorts of bias that does exist.

Because of the small numbers involved in this aspect of the work we would not wish to take our arguments too far based on this and other analyses of the data, but we can go some way to increase our confidence in what has been concluded earlier, where we have related our results, on the whole, to the final step in the identification process and the teachers that were involved, rather than to the 'system' from which the list evolved.

TABLE 1

MAIN VARIABLES THAT CHARACTERISED SELECTION OF THE TOP 10% LISTS FROM THE TEACHERS TOP 20% LISTS (ROUND 2, YEAR 1)

SCHOOL	ENG	ρ	n	FR	ρ	n	PHYS	ρ	n	MATHS	ρ	n
01	A	0.47	35	NT	0.48	26	G	0.37	36	F	0.47	25
				SP	0.46	26						
02	-	-	-	LU	0.42	42	B	0.51	40	VN	0.56	46
				SP	0.38	42	C	0.50	40	AR	0.52	46
							E	0.44	40	CSA	0.54	46
03	VR	0.63	22	D	0.41	24	B	0.66	24	SC	-0.35	30
04	I	0.27	37	SP	0.54	35	R	0.30	40	C	0.52	62
										NA	0.53	62
10	VN	0.59	22	NT	0.44	22	SEX	0.68	22	LU	0.59	32
	LU	0.65	22									
11	CSA	0.42	23	-	-	-	SR	0.56	23	AR	0.52	34
12	FLU	-0.45	22	NA	0.48	24	E	0.34	25	I	0.58	24

Pearson Correlations of background test with selection (category score for pupils selected on top 10%, 0 for pupils not selected) are shown together with numbers in the sample for each correlation.

APPENDIX 7

The Mathematics checklist

Validation

We have discussed, in early chapters of this report, the diffiulty of finding the appropriate aptitudes for a checklist which is to be used in the classroom. The items must relate to what is observable as well as being valid for the subject in question. We have seen that there is some uncertainty concerning items that appear on many checklists which refer more to general than specific ability.

In searching for valid checklist items for mathematics these problems were to be heeded. There have been many research initiatives to identify the characteristics of able mathematicians in particular aspects of the subject. Mathematics involves problem solving of a particular kind and there has been considerable research both in problem solving of a general kind and problem solving in various branches of mathematics[1]. These researches are generally concerned with testing a model or strategy that is employed in problem solving. From these studies, however, characteristics of able problem solvers emerge. Kulm and Bussmann (1980) in evolving arationale for their own 'Phase-Ability Model' list many of the characteristics that previous researches have discovered. One notable contribution comes from the work of V.A. Krutetskii (1976). Characteristics that he has listed as valid indicators of mathematical ability are much the same as other researchers have discovered independently in a variety of studies.

Krutetskii's work has been accepted widely both by researchers and classroom practitioners. In particular it is the only reference given for the identification of mathematical ability in the Cockroft Report ('Mathematics Counts', 1982) which has been the basis of much planning for mathematics teaching in Britain. It is also evidenced by its inclusion as a main reference for those concerned with the identification of the able mathematicians in the Schools Council Booklet (Straker, 1981). It represents the best starting point from mathematical problem solving research that is available for the compilation of a checklist for classroom use. Krutetskii took as his starting points what was already kown of the characteristics of able children from the international literature, what mathematics teachers considered the important qualities and what characterised mathematicians who had achieved eminence. He used as his model of the problem solving process one involving the three stages: gathering information, processing information and information retention. He identified three types of mathematical mind: The Analytic who thinks in verbal, logical terms, the Geometric who thinks in visual, pictorial terms and the Harmonic who combine both of the other two characteristics.

By constructing ingenious challenges of particular kinds and evaluating the responses of pupils of differing mathematical ability, Krutetskii was able to isolate the following characteristics:-

1. Obtaining mathematical information
   - A. The ability for formalized perception of mathematical material, for grasping the formal structure of a problem.

2. Processing mathematical information
   - A. The ability for logical thought in the sphere of quantitative and spatial relationships, number and letter symbols; the ability to think in mathematical symbols.
   - B. The ability for rapid and broad generalization of mathematical objects, relations, and operations.
   - C. The ability to curtail the process of mathematical reasoning and the system of corresponding operations; the ability to think in curtailed structures.
   - D. Flexibility of mental processes in mathematical activity
   - E. Striving for clarity, simplicity, economy, and rationality of solutions.
   - F. The ability for rapid and free reconstruction of the direction of a mental process, switching from a direct to a reverse train of thought (reversibility of the mental process in mathematical reasoning).

3. Retaining mathematical information
   - A. Mathematical memory (generalized memory for mathematical relationships, type characteristics, schemes of arguments and proofs, methods of problem-solving, and principles of approach).

4. General synthetic component
   - A. Mathematical cast of mind.

If it is accepted that these characteristics are appropriate they can form the basis of the checklist for teachers, though as we shall see further validity work was thought to be necessary for our own particular application. It must be stressed that these are items referring to aptitude and it could be argued that attitude and motivation should be important components of the identification process. This will be considered elsewhere, but in the formation of the checklist, which is concerned with the identification of ability alone, leaving the exploitation of this ability through motivation as an important but secondary matter, only these items have been considered.

At this stage the main concern was that which was discussed earlier, of making the checklist into a form where the items related to what might be easily understood and observable by the teacher. This was achieved to some extent by looking closely at the sorts of challenge that Krutetskii gave and deriving classroom observable items related to them. The list included, therefore, several titems which related to the same mental ability tha Krutetskii had investigated. For example items 11 and 12 in the final form of the checklist: trying alternative approaches mid-way in a proof without losing way and finding alternative methods of solution are both measures of the ability 2D, flexibility of the mental processes. The teacher is much more able to observe one or both of the checklist items than to be sensitive to clues relating to the single flexibility item.

The complete relationship between the checklist items and Krutetskii's aptitudes is:

Krutetskii's Aptitude	Relevant Checklist Items
1 A	1, 2, 3
2 A	4, 5, 6, 21, 22
2 B	7, 8
2 C	9, 10
2 D	11, 12, 19
2 E	13, 14, 15, 20
2 F	16
3 A	17, 18
4	22.

Validation Rationale

The importance of Construct Validity in relation to psychological measurements was established in the 1950's. Cronbach and Meehl (1955) suggest that it should be incorporated in test design whenever the researcher needs to establish which constructs (that is, human qualities or attributes) will account for the variance in test performance. Thurstone (1952) highlighted the important aspects of the situation. An accepted method of validating a test had been an outside criterion. Commonly this would be based on what could be reasoned as reliable judgements by experts. Thurstone criticised the correlation method as being 'too coarse' and also doubted the wisdom of the criterion based on 'judgements' of this kind. The main point is that a method which concentrates on validation of the test after its construction is inferior to one that concentrates the effort on the construction itself. Thus, while accepting that the test in its final form should be externally valid, the researcher can take steps in the test design to ensure that internal validity is present by establishing that it contains appropriate constructs.

The checklist of aptitudes, upon which the ability of pupils in mathematics is to be assessed is, indirectly, a test of mathematics for these pupils, so that the items on the checklist should be validated as if they were test items. It must be established that these do indeed represent constructs that are appropriate for the judgement of ability in mathematics. This validity can to some extent be assumed, in that Krutetskii's (1976) work has established the independent components of mathematical ability from which the checklist has, on the whole, been derived. Furthermore the checklist items are quite clearly applicable to mathematical activity - what we mean by mathematics. We are, however, applying the observations - the test - to a particular age group of children in a particular educational system so that we need to establish whether these checklist items are relevant to that situation. We also need to check if they are the items that will show meaningful differences in children's performance at the top of the mathematical ability range. Krutetskii's work can only suggest that this is possible. His focus was on the components of mathematical ability across a broad band of pupils and, though the age group that their own research examined was included in his work, so were children of quite different ages.

Chronbach and Meehl (1955) suggest various strategies for checking construct validity among which is 'group differences'. This has been adopted for the present study. The method was to compare a group of pupils of high ability in mathematics with a group of pupils of lower ability on a series of mathematical challenges. These challenges were designed to show their ability levels on each of the checklist aptitudes.

We will now describe the groups of pupils who were used for this comparison.

Selection of pupils for the validity check

In its validated form the checklist would be used in the classroom to classify children relative to one another, according to clues that they would give to aptitudes defined by the checklist items. It would be necessary for this classification to enable those of high relative ability to be distinguished not just from those of low ability but also from those of moderately high ability. This is because Mathematics is commonly taught in ability sets so that, approximately, the top 20% for a school year group may be in one teaching set. The task would be to identify subgroups such as the top 10%, top 5% or top 1% from within the top 20%. The checklist must, then, contain items that have been shown to be valid components of mathematical ability and also represent constructs where recognisable differences occur within the top percentiles of the ability range. This would suggest that the appropriate way to conduct the validity check would be to select a group of very able mathematics pupils and compare them with a set of moderately able pupils rather than those of very low ability. A second important reason for this strategy is that the further apart our groups are on mathematical ability, the more likely it is that they will differ on abilities other than mathematical ones. Thus invalid items could be kept on the checklist in error. The following analysis gives evidence for this hypothesis.

Appendix 1 gives details of how a formula for prediction of O-level success was derived by the technique of Multiple Regression from the O-level scores (at age 16) and scores (at age 13) on a variety of tests and questionnaires. The appendix also lists the tests that were administered with the abbreviations that will be used here. These same tests were administered to other third form (13 year old) pupils and predicted O-level scores were obtained. These predicted scores represent a measure of mathematical potential for O-level.

Table 1 gives the means and standard deviations for the group of children whose scores were used to produce the prediction equations, with a correlation coefficient that was computed between each variable and mathematics O-level scores[2].

TABLE 1

STATISTICS CONCERNING 3rd FORM TEST SCORES AND O-LEVEL RESULTS ON A GROUP OF 131 PUPILS

VARIABLE	MEAN	STANDARD DEV.	NO. OF CASES	PEARSON CORRELATION WITH O-LEVEL SCORE
FLU	47.0	11.3	131	0.09
FLEX	34.6	6.9	131	0.10
ORG	21.2	8.9	131	0.05
A	4.2	1.3	131	-0.11
B	3.1	1.4	131	0.02
C	6.3	1.5	131	0.15
D	7.2	1.2	131	-0.04
E	3.4	1.0	131	-0.15
F	3.1	1.2	131	-0.02
G	7.7	1.5	131	-0.06
I	3.5	1.2	131	0.00
VR	31.9	7.5	131	0.35
NA	18.5	4.6	131	0.47
AR	38.6	4.7	131	0.33
CSA	42.0	9.2	131	-0.13
MR	47.6	8.4	131	0.34
SR	36.4	10.4	131	0.33
SP	66.3	15.4	131	0.02
LU	34.7	8.1	131	0.19

These figures are based on a resticted sample made up of only those who were considered able to achieve a grade C or above at O-level. The prediction equations have been used over the whole range of ability. Correction of the correlation coefficient has been applied according to the method described by (among others), Guilford and Fruchter (1978). This is based on the figures in Table 1 and standard deviations derived from the sample of pupils drawn from the full ability range on whom the prediction equations have been used.

TABLE 2

SHOWING CORRELATION COEFFICIENTS THAT HAVE BEEN CORRECTED FOR RESTRICTION OF RANGE, BETWEEN THE BACKGROUND TEST SCORES AND MARK AT O-LEVEL MATHEMATICS

VARIABLE	STANDARD DEVIATION[3]	CORRECTED CORRELATION WITH O-LEVEL SCORE
FLU	11.8	.10
FLEX	7.4	.11
ORIG	10.1	.06
A	1.4	-.12
B	1.6	.03
C	1.8	.17
D	1.8	-.06
E	1.2	-.18
F	1.3	-.02
G	2.0	-.08
I	1.3	.00
VR	8.8	.40
NA	6.4	.58
AR	9.6	.58
CSA	11.9	-.19
MR	9.7	.39
SR	10.6	.34
SP	15.7	.02
LU	9.2	.22

The variables from this list that contributed significantly to the variance in O-level performance from regression analysis were NA, MR, Ar, E, VR[4]. The variable SR which is correlated .34 with O-level mark, did not contribute to the variance[5] in addition to these because of the intercorrelations (.36 with NA, .42 with Ar, .53 with MR) with other variables, though it must still be considered, in isolation, as an indicator of potential in mathematics. LU is a weaker indicator in the same way.

A new sample of pupils whose predicted O-level scores had been calculated were divided into 10% bands based on these predicted scores. One way analysis of variance was then carried out on each of the variables. In this way it was possible to see which of them gave significantly large differences between the top 10% and the lowest ability bands of maths ability and which gave also significant differences between the top 10% and higher ability bands (particularly the band of pupils between the 80th and 90th percentile on predivted O-level score).

The following table summarises these results, withathe following coding for the percentile ranges:

GROUP 1 : (90-100)
GROUP 2 : (80-90)
GROUP 3 : (70-80)
. .
. .
. .
GROUP 10 : (0-10)

TABLE 3

SHOWING WHICH TESTS GAVE SIGINIFICANT DIFFERENCES FROM THE ANALYSIS OF VARIANCE APPLIED TO 1870 PUPILS IN 10%BANDS BASED ON PREDICTED MATHS O-LEVEL SCORE

TEST	F PROBABILITY FROM ANALYSIS OF VARIANCE	GROUPS WHOSE MEAN TEST VALUE GAVE SIGNIFICANT T VALUES COMPARED WITH THE MEAN FOR GROUP 1	
		(T. Prob. .05)	(T. Prob. .01)
Neatness	.0000	7 - 10	6 - 10
FLU	.0000	3 - 10	4 - 10
FLEX	.0000	2 - 10	3 - 10
ORIG	.0000	2 - 10	3 - 10
Social Class	.0000	2 - 10	2 - 10
A	.0000	4 - 10	5 - 10
B	.0000	3 - 10	4 - 10
C	.0000	2 - 10	2 - 10
D	.1712	8 - 10	8 - 10
E	.7152	-	-
F	.1283	8,9	-
G	.0000	4 - 10	4, 6 - 10
I	.0003	4 - 10	4 - 10
VR	.0000	2 - 10	2 - 10
NA	.0000	2 - 10	2 - 10
AR	.0000	2 - 10	2 - 10
CSA	.0000	3 - 10	3 - 10
MR	.0000	2 - 10	2 - 10
SR	.0000	2 - 10	2 - 10
SP	.0000	2 - 10	2 - 10
LU	.0000	2 - 10	2 - 10

If groups of able mathematicians had been compared with groups of considerably less able mathematicians (based on the criterion of O-level potential) in order to see on which of these tests there was also a difference there would have been a number of misleading results. For example if pupils in the top 10% had been compared with any groups below the 30th percentile, taking 5% significance as sufficient, high scores on all of the variables except Attitude E would have appeared to be associated with mathematical ability. If the top 10% had been compared with the next 10% (group 2) only the variables FLEX, ORIG, Social Class, C, VR, NA, AR, SR, SP and LU would have shown the association. If 1% significance is taken as the appropriate level for showing a real difference between groups, the picture does not change appreciably for comparisons with the lowest percentiles, but in the Group 1 and Group 2 comparison the identification would be that Social Class, VR, NA, AR, MR, SR, SP and LU were the variables associated in some way with mathematical potential. This is close to the indications from the results of the regression analysis reported earlier. This gives confirmatory evidence that in order to validate the checklist items a way of comparing two groups, one of high ability and the other of moderately high ability, should be found, so that the spurious differences which would arise between groups of wider mathematical ability can be avoided, maximising the possibility of finding items that truly represent mathematical aptitudes.

The question, then, was how far apart on mathematical ability our two groups should be. It was not possible to have groups as close in ability as would be generally found in a top school's set because of our inability to discover a background criterion that would form the basis of selection process for the two groups. The risk involved in selecting pupils in groups further apart in ability would be that differences on the checklist items might be observable in these children but not observable in the classroom where the checklist would be used to recognise smaller differences between pupils. We saw in chapter twelve, however, that the differences were in fact recognisable between such children.

Selection of the highly able mathematics pupils was done with care in that the criterion for inclusion in the group had to be different from that which was being examined by the research. Thus teacher nomination was necessary though insufficient. Pupils who would be in the classes which were later to be observed were not used for this purpose. A search was made through England for any children who for reasons other than teacher nomination or normal school based testing had shown themselves outstanding in mathematics. Contact was made with organisers of the main provision programmes[6] for pupils in mathematics with a request for their co-operation regarding their most able pupils in the 13 to 14 year age group who were undoubtedly able at mathematics, this to be justified by some valid criterion. Descriptions of the pupils is given in Table 8 and the independent criterion stated. As final confirmation of this in all cases, the children's ability was also given unreserved support by school performance and by their mathematics teachers. They each can be considered as having emerged from a population of several thousand by vitue of their high ability. For example if a pupil was included by virtue of performance in a provision programme in mathematics, there would have been a number of screening procedures in order that he/she was included in the scheme and then selection would be as a result of performance in the high level challenges within the provision programme and in comparison with other able pupils.

TABLE 4

DESCRIPTIONS OF PUPILS OF HIGH MATHEMATICAL ABILITY WHO WERE USED IN THE VALIDATION OF THE CHECKLIST

Source[7]	Age at time of validation work
1. Joint third in Mathematics Olympiad organised by local Education Authority (1st and 2nd places were taken by much more experienced 17 year olds and many other older and very able children performed less well in this competition). Nominated by County Adviser who organised the competition and who is also an expert in the field of mathematical giftedness.	14
2. As 1: the other joint third prizewinner in the LEA Mathematics Olympiad.	15
3. Unable to compete in the same competition as 1 and 2 due to ill health, but normally surpasses mathematical success of 2 both in school and in County Provision programme for gifted mathematicians.	14
4. Identified within a provision scheme for gifted mathematicians set up by the Mathematics Department of a university in co-operation with local schools.	13
5. From the same provision programme as 4.	12
6. Identified within a major County-wide provision programme for gifted children.	13
7. From the same provision programme as 6.	14
8. Identified from performance in mathematics masterclasses for secondary school pupils.	14
9. From the same masterclasses as 8.	14
10. Identified by primary school as gifted in mathematics. Given special provision-supported by LEA. Achieved Grade A at O-level at age 12.	13
11. Highest of approximately 4,000 pupils who formed a represetntative sample from one LEA, tested on a combination[8] of tests used for prediction of O-level performance.	13
12. Second highest on the same testing procedure as 11.	13

Selection of the pupils of moderate ability was not such a problem. It was possible to find 13 year old pupils

a) who were in the top set in their year group;

b) whom their teacher confirmed as in the bottom half of this set on general performance;

c) whose predicted O-level score was above that appropriate to a grade C but below that appropriate to a top A grade.

Twelve such pupils were selected. The predicted scores and estimated grades are shown in Table 5.

TABLE 5

PREDICTED O-LEVEL SCORES AND GRADES FOR THE PUPILS OF MODERATE ABILITY WHO WERE USED IN THE VALIDATION WORK

CODE	MARK	GRADE
MM1	179.06	Low A
MM2	166.40	High B
MM3	144.10	B
MM4	142.76	B
MM5	142.12	B
MM6	138.15	Low B
MM7	167.87	High B
MM8	157.95	B
MM9	153.72	B
MM10	149.67	B
MM11	139.14	Low B
MM12	136.31	Low B

The approximate grade ranges are:

170+	A
130-169	B
105-129	C

Validation technique

A number of possible strategies could have been tried in order to assess each of the pupils on each item. An efficient method in terms of administration would be to construct mathematics test papers either containing questions with specific or general mathematical challenge and attempt to isolate strengths and weaknesses of the pupils on the particular aptitudes in the checklist. This method, whether the test was constucted particularly for this purpose or whether use was made of tests available from publishers, suffers from the difficulty that assumptions have to be made about the sorts of mathematics topics that would be suitable, and what might constitute a particular sort of challenge. Inevitably the test becomes a test of attainment rather than aptitude. Indeed, a search through published psychological tests as wide as those reviewed in the Seventh Mental Measurements Yearbook[9] and ones of particular relevance to Great Britain such as those published by the National Foundation for Educational Research[10] reveals only tests of general mathematical attainment and any tests in specific mathematics areas are in broad categories such as arithmetic or geometry.

The checklist has evolved from the work of Krutetskii and he used the technique of setting individual children mathematical challenges and remaining with them to monitor their responses. He compiled a variety of mathematical challenges, each applicable to a particular mathematical aptitude. It would seem appropriate to use his bank of questions in a similar way in order to check the validity of the items in the context of this research. This, however, is not appropriate for two reasons. Firstly he chose his problems for the particular pupils with whom he was working, based on their stage of mathematical growth in terms of topics covered and developmental stage concerning a particular aptitude. Secondly the number of problems that he accumulated would take far too long to administer to an individual. Different children were used in different stages of his work over the years that it was done. The method that he used, however, is appropriate for this work and many of his ideas for mathematical challenges can be exploited directly or developed for the purpose of this research.

Other researchers such as Kulm (1980) who wish to gain insights into the mental processes of problem solving argue the case for working with individual pupils. For the purpose of this research there are a number of strengths of the method.

a) Pupils can be set mathematical tasks that are appropriate to their level of attainment. For example a rapid transfer from $(a + b)^2 = a^2 + 2ab + b^2$ to an expansion of $(E + F - G)(E - G + F)$ would indicate a strong aptitude for generalisation for a child whose algebra experiences were limited, but give no valid clue at all in the case of a child who had practiced this kind of thing. A child could be examined twice over a relatively short interval of time on a challenge such as this and give conflicting clues to his aptitude because of work done in the intervening time. The interaction that is possible with examiner and pupil in an interview situation would enable a more reliable interpretation to an appropriate challenge to be made.

b) The feedback that a pupil gives to the interviewer as he is asked to work aloud will enable the interviewer to clear up trivial misunderstandings of what is expected of the pupil, to overcome ambiguities, to break a question down into smaller steps or ask slightly different questions that will give the pupils the best opportunity to show his/her levels of ability.

c) The interviewer can vary the questions and ordering of questions to suit the particular pupils and clues that he/she is giving so that opportunities for deeper insights can be used as they occur.

d) By being asked to 'think aloud' on certain kinds of challenge the stages and patterns of reasoning can be monitored more effectively. More clues with greater possibilities for interpretation are available than if the pupil has been given a set of tasks to answer alone and on paper. The opportunities for an examiner in the latter situation are resticted by what the pupil can communicate by writing.

e) Some aptitudes such as the ability to show flexibility of thought in solving mathematical problems would present large difficulty to the constructor of a pen and paper test, yet it can be detected by an interviewer who works with a pupil who is 'thinking aloud'.

The method, then, is for an interviewer to have at his disposal a variety of mathematical challenges, each designed to monitor a particular mathematics aptitude from the checklist - sometimes more than one aptitude: indeed it often turns out that a challenge that an interviewer initially designed for a certain aptitude reveals something quite unexpected on a different aptitude. The examiner must be sensitive to the possibilities for a particular pupil and he can think of himself as giving opportunities for the pupil to display his/her highest possible level on a particular aptitude. The interviewer must, of course, beware of the dangers, in the interaction of leading the pupil in such a way that what the pupil achieves is more as a result of this leading than his/her own aptitude. This problem is minimised when the interviewer adopts the stance of follower through a path of logic rather than leader, asking questions rather than suggesting answers and then only so that the dialogue will continue. The success depends on the sensitivity of the interviewer to what is achieved as a result of the pupil's own ability and what was achieved otherwise.

As far as possible pupils who have achieved similar points in their mathematical growth should be given similar challenges so that comparisons are direct. This was found to be possible in this particular validation work on a considerable number of occasions, mainly because pupils were in the same school year groups and had covered similar topics.

The choice of problems for inclusion in the 'pool' available to the interviewer took considerable effort. For example, the problems that were developed by the A.P.U. were considered, but it was found difficult to categorise them in the manner suitable for the mathematical profile that was to be built up. Furthermore the problems were designed for assessment over a wider range of overall mathematic ability in children. The problems that were categorised as the ones which only the ablest pupils could solve were noted and were available when final selections were made. Another useful source for establishing initial norms was the NFER Secondary Mathematics Item Bank (Purusbotheman, 1976), but here again, though the questions were graded according to difficulty and also categorised, the categorisation did not fit easily into the profile that was to be validated. The most appropriate method was to use the ideas of Krutetskii as far as possible, modifying where necessary, and then seeking problems from a variety of sources to challenge other of the aptitudes on the checklists. The sources are listed along with the problems at the end of this appendix.

Discussion of the findings

Later in this section we have listed the challenges that were available as a resource 'bank' from which appropriate challenges could be chosen. Most of these were prepared in advance of the first testing session but several were added to the 'bank' as they were needed for a particular pupil, to maximise the opportunity for him/her of showing aptitude levels with challenges appropriate to his/her expertise. Krutestskii's 26 series of problems have been the inspiration for many of the challenges and some are direct copies from his work. Even with his large list there remain some particular challenges that were appropriate to individual pupils and were introduced to the testing session without previous preparation. For example one of the able pupils showed a particular concern that he needed calculus for a problem that was worrying him. The examiner proceeded to explain some elementary principles and the dialogue continued so that in the space a twenty minutes a month's sixth form calculus was covered! The opportunity to gain insights into the pupil's capabilities for generalisation, following proofs and unwillingness to put an interesting problem aside was thus exploited. It would be impracticable to list every instance of this kind, though notes have been kept from each testing session concerning every aspect of the work covered.

It is worth noting too that a challenge designed to test a particular aptitude did not always have the expected results. For example in the series designed to see how freely a pupil gave ideas to solve a problem, a particular pupil refused to give the problem back until it was solved, giving many extra unexpected clues to other aptitudes, particularly item 20 on the checklist.

We have included detailed descriptions of two complete sessions, one with a pupil of high ability and the other with a pupil of moderate ability. In general it was arranged to have available 2½ to 3½ hours per pupil, allowing time for break and refreshment as necessary. This was found quite adequate for all purposes though in the case of many of the moderately able pupils they reached limits on all challenges in less than 2½ hours and some of the ablest pupils went on for a little over 3 hours.

The observed aptitude levels on each of the checklist items were categorised as follows. Three levels were possible (L (low), M (medium), H (high)). These levels evolved quite clearly from the responses of the pupils rather than from preconceived ideas. There was little doubt about which level was applicable in the final analysis.

ITEM 1

L. Fails consistently on all but the simplest problems from series 1a, 1b and 1c. When confronted with problems or puzzles takes considerable time to understand the information. Fails to get started on the problems that involve the more complex information.

M. Succeeds on all but the hardest questions on at least one of series 1a, 1b, 1c, but fails consistently on a significant number of questions. Is able to cope with information of a moderately complex nature but takes considerable time to understand the more complex ones, often missing and not using some of the facts.

H. Succeeds without hesitation on all of the problems from 1a, 1b, 1c. Shows confidence in understanding the relationships between items of information in the most complex problem and shows due regard to these relationships in the solutions to the problems. (The examiner was careful to note if there was any differences in reaction to problems of the spatial or analytic kind to monitor Items 22a and b.)

ITEMS 2/3

L. Fails consistently on all but the simplest examples from series 2. Gets confused in problems with redundant or superfluous information.

M. Succeeds in all but the hardest questions from series 2, but fails consistently on a significant number of these questions. Displays lack of confidence with some other questions involving irrelevant information which becomes an obstacle to solving the problem.

H. Succeeds without hesitation on all of the problems from series 2. Is able to see what is relevant in other problems with superfluous information, rejecting irrelevant data immediately.

ITEM 4

L. Is unable to give any ideas when confronted with an unfamiliar problem.

M. Can give ideas that have some relevance, but many of them would not lead to a successful solution. Many of the more plausible ideas show no depth of insight and the pupil cannot develop them to the next stage of

225

the solution.

H. In cases where weaker pupils fail can give ideas that can be developed into a solution. Shows the ability to develop them himself (herself) and that the ideas come from insights into an overall plan for the solution. Avoids guesswork.

ITEM 5

L. Every step in a solution must be carefully explained. Loses track of the logic in a proof that becomes complex or that involves more than a few steps.

M. Able to follow linear paths of logic but becomes confused when the proofs become complex, has difficulty in overcoming the difficulty.

H. Able to follow complex proofs with ease. Rarely gets confused and is able to overcome the confusion after further attempts.

ITEM 6

L. Is only able to do the kind of problem for which there is a precise algorithm that has been adequately demonstrated.

M. Is able to find solutions to problems which require some variation from a known algorithm, but displays considerable weakness in problems that are quite different from the types that have been tackled in familiar mathematics topics.

H. Shows the ability to solve problems of a new or unusual nature.

ITEM 7

L. Cannot give convincing suggestions concerning the links between mathematics topics. Fails to make use of knowledge from one topic to solve problems in another.

M. Gives some suggestions about how one topic links with others but cannot readily make use of the links to solve problems, missing opportunities that more superior pupils exploit immediately.

H. Describes deep relationships between maths topics. Displays an ability to readily make use of one topic to solve problems in another in such a way that the pupil is showing a deep sense of the pattern of relationships.

ITEM 8

L. Unable to describe a method generally, or can only demonstrate specific examples. Is unable to follow or produce generalised proofs. Fails on the problems testing ability to generalise.

M. Has difficulty in describing a general method without examples to illustrate the strategy. Is able to follow or produce only simple algerbraic proofs. Succeeds finally, but only after careful guidance on the problems concerned with generalisation.

H. Readily describes a method in general terms. Confident with algebraic solutions. Succeeds with ease in the generalisation tests.

ITEMS 9/10

L. Cannot follow a proof unless every step is included. Becomes confused when even one step is assumed.

M. Is able to cope with very straightforward jumps of logic but does not display the ability to make logical jumps in his/her own proofs. Becomes confused in cases where large steps are assumed.

H. Not only follows proofs when large amounts of logic are assumed but displays a strong tendency to make these valid logical jumps himself/herself with the ability to fill in the steps if required.

ITEM 11

L. If one method fails in a proof becomes disorientated unless a fresh start is made. Tends to give up rather than try again.

M. Is able to follow someone else's alternative strategy mid-proof providing it has not developed very far, but does not show any initiative to seek an alternative path alone.

H. Can easily restart a proof from any valid point and does not become disorientated even in complex situations. Shows initiative in trying alternative methods from appropriate points in the proof when one part of the argument develops in an unsatisfactory way.

ITEM 12

L. Cannot find alternatives when one method fails. Accepts the first solution that is achieved.

M. Keeps trying, but does not find it easy to establish a variety of approaches.

H. Freely gives more than one method where possible.

ITEM 13

L. Neither seeks an elegant solution nor shows any criticism of an inelegant one. Any answer will do.

M. Is prepared to try alternative approaches based on work that is understood but stops short of seeking the most elegant if it requires an extra effort with unfamiliar work.

H. Is constantly seeking to improve solutions, showing delight in the most elegant and will tackle new topics in order to achieve a satisfactory solution or an understanding of someone else's proof.

ITEM 14

L. Does not notice when even trivial slips of logic are made. Cannot correct an invalid argument. Even prepared to accept false conclusions, suggesting that mathematics can give wrong answers.

M. Notices trivial slips of logic, but fails to notice subtle illogical steps.

H. Quick to see when a step is illogical. Will not accept an illogical conclusion, but searches for the invalid steps in the manner of solving a mathematical problem.

ITEM 15

L. Makes little effort to seek understanding even in simple proofs where a particular step was not clear.

M. Seeks to understand any of the current work in the way that it has been taught, but shows no inclination to gain understanding outside the explanations given with the present topic.

H. Makes determined efforts to understand all that has not been adequately explained, studying topics outside the syllabus if necessary. Has an insatiable appetite for knowledge and understanding.

ITEM 16

L. Is completely muddled when a 'reverse' question follows a 'direct' one.

M. Takes time to adjust to a 'reverse' problem, in order that the mind can forget the 'direct' problem and take on the 'reverse' problem as a new problem in a 'direct' way.

H. Adjusts immediately from the 'direct' to the 'reverse' problem.

ITEM 17

L. Remembers how to tackle a problem by learning every step of an algorithm and can only solve the sorts of problem that have been practised.

M. Remembers a method by its algorithm but is able to tackle problems requiring adaptations of the algorithm.

H. Remembers proofs and methods by the essential details and general methodology rather than every step of an algorithm.

ITEM 18[11]

L. Possesses only short term memory of the current work. Has to relearn rather than revise.

M. Poor long term memory, but can regain competance by revision.

H. Rarely needs to revise a topic. The essential details are remembered over long periods of time.

ITEM 19

L. Becomes tired quickly, even on routine work.

M. Shows no fatigue on familiar work but tires when intense concentration or logical arguments are needed.

H. Does not tire even after long periods of intense mathematical activity.

ITEM 20

L. Gives up when a problem is not solved easily.

M. Will keep trying to solve a problem in lesson time but soon forgets it after the lesson is over. Is prepared to ask for help rather than overcome difficulties.

H. Will not leave a problem unsolved even outside lessons. Shows great desire to find his/her own solution rather than seek help. Becomes pre-occupied with the problem at the expense of other work or pleasure until it is solved.

ITEM 21[11]

L. Makes frequent errors even on simple computation.

M. Competant with any arithmetic but needs to use paper and pencil.

H. Very quick and accurate both with pencil and paper and with mental arithmetic.

ITEM 22[11]

A. Tends to solve problems in a spatial manner even if an easier analytic solution is available

B. Shows strength in analytic/logical work, but weakness in spatial work.

C. Is able to solve problems of both ANALYTIC and SPACIAL kinds showing no particular bias.

The case studies will illustrate how these levels have been applied. The following table summarises the findings, firstly from the moderately able group and then from the pupils of known high ability.

TABLE 6

CHECKLIST LEVELS FOR MODERATELY ABLE PUPILS

Pupil Code	Checklist item 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
MM1	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	-	M	M	-	B
MM2	M	M	M	M	M	M	M	L	M	M	M	M	M	M	M	M	M	-	M	M	-	-
MM3	M	M	M	L	L	M	L	M	L	L	L	L	L	L	L	L	L	-	-	L	-	B
MM4	L	L	L	L	M	L	M	L	L	L	L	L	L	L	L	L	L	-	L	L	-	B
MM5	M	M	M	M	M	M	M	L	M	L	M	M	M	M	M	L	L	-	M	M	-	-
MM6	L	L	L	M	M	L	L	L	L	L	M	L	L	L	L	L	L	-	L	L	-	-
MM7	M	L	L	L	M	M	L	M	L	L	L	L	L	M	L	L	L	-	L	L	-	-
MM8	M	L	L	L	L	L	M	L	L	L	L	L	L	L	L	L	M	-	L	L	-	-
MM9	M	M	M	M	M	M	M	M	M	L	M	M	M	M	M	L	M	-	M	M	-	-
MM10	M	M	M	H	M	H	M	M	M	L	L	H	M	L	M	L	M	-	M	H	-	-
MM11	M	L	L	L	L	M	M	M	L	L	L	L	L	L	L	M	L	-	L	L	-	-
MM12	L	L	L	M	L	M	L	L	L	L	L	L	L	L	L	L	L	-	L	L	-	-
Totals																						
H	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	-	0	1	-	-
M	9	6	6	6	8	8	8	6	5	2	5	4	5	5	5	3	5	-	5	4	-	-
L	3	6	6	5	4	3	4	6	7	10	7	7	7	7	7	9	7	-	6	7	-	-

TABLE 7

CHECKLIST LEVELS FOR PUPILS OF KNOWN HIGH ABILITY

Pupil Code	Checklist item 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
HM1	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	H	M	-	-
HM2	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	H	H	H	-
HM3	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	H	M	-	-
HM4	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	H	H	-	-
HM5	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	B
HM6	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-
HM7	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	H	H	-	-
HM8	H	M	M	H	H	H	H	H	M	M	H	H	H	H	H	H	H	M	H	M	-	-
HM9	H	H	M	H	H	H	H	H	H	H	H	H	H	H	H	H	M	H	H	M	-	B
HM10	H	H	M	H	H	H	H	H	H	H	H	H	H	M	H	H	H	H	H	H	-	-
HM11	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	B
HM12	H	H	H	H	H	H	H	H	H	H	H	M	H	H	H	H	H	H	H	H	H	B
Totals																						
H	12	:	9	12	12	12	12	12	11	11	12	11	12	11	12	12	11	6	12	8	3	-
M	0	1	3	0	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	4	0	-
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-

It is readily apparent that significant differences were found on most of the checklist items. To confirm what seems obvious from the figures the Gamma test was applied to confirm that high ability on each of these checklist items can be associated with high mathematical ability.

TABLE 8

ASSOCIATION OF CHECKLIST APTITUDE WITH MATHEMATICS ABILITY

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20
GAMMA	1	1	1	1	1	1	1	1	1	1	1	.985	1	1	1	1	1	1	.915
SIG. OF GAMMA	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****

**
* indicates that we would obtain this level of Gamma with a probability of less than 0.1% if there was no association. The maximum value of Gamma is 1.

The checklist items on which it was not possible to draw confident conclusions were Items 18, 21 and 22.

The testing session was not designed to test in any depth the pupils' long term memory, though various aspects of the work concerned topics that the pupils had previously covered in school. In the case of the pupils of moderate ability, the clues gained were not considered adequate to make proper judgement. In the case of those of high ability adequate clues were obtained from seven of the twelve. The main item of memory was item 17 which concerned the way in which pupils remembered mathematical material. On this aptitude the rpupils of high mathematical ability demonstrated the ability to memorise essential details and general strategies more strongly than those of lesser ability.. It could be reasoned that this would enable them to remember mathematical methods and proofs more readily over the long term. One could further reason that a pupil is unlikely to make progress with his mathematics without a satisfactory memory. Too much time would be spent relearning old work at the expense of the new. This reasoning and the evidence gained from the seven highly rated pupils would indicate that this item should be kept on the checklist, being more appropriate in a class situation than an individual test environment.

There was no specific test of arithmetic included in the testing sessions, so that clues to ability were limited to the computations that naturally occurred. This was considered inadequate for proper judgement of all but three of the pupils of high mathematical ability, who gave clear evidence of their high abilities at mental arithmetic. Krutetskii (1976) considered the situation regarding arithmetic skills and quoted examples where eminent professors of mathematics were very poor at computation. It could be reasoned that an able pupil might consider arithmetic skills as rather unimportant compared with the development of logical abilities and not give them overmuch attention, whereas arithmetic is an important aspect of work in primary schools and early secondary school so that those of lesser overall mathematical ability could train themselves to a high level and thus obtain high rankings in the early school years. The algorithmic nature of much work in arithmetic would allow even a pupil with a narrow range of mathematical aptitudes to succeed. It would seem that the understanding of number and of arithmetical operation is a higher level challenge than that of simple computation, but one that cannot be assessed from achievement tests. Table 9 lists the percentile scores that were obtained by the moderately able pupils on the Numerical Ability test in the DAT battery.

TABLE 9

NATIONAL NORM PERCENTAGES ON DAT NUMERICAL ABILITY TEST FOR THE TWELVE PUPILS OF MODERATE MATHEMATICS ABILITY

PUPIL CODE	PERCENTILE
MM1	99
MM2	99
MM3	95
MM4	95
MM5	90
MM6	95
MM7	97
MM8	95
MM9	90
MM10	90
MM11	90
MM12	50

All but one of the pupils has shown attainment above the 90th percentile on national norms for numerical ability and seven of the twelve above the 95th percentile. This provides evidence that pupils of moderate overall mathematical ability can be very proficient with numerical work.

The checklist item was retained, therefore, with reservation. Judgement of arithmetical competence was to be used in the identification strategy in the sense that understanding of numerical relationships was important. Achievement both in mental and paper and pencil computational skills was to be of secondary importance. It was also considered that further insights into the relevance of arithmetic to overall ability could be gained by its inclusion on the list.

It was possible in a total of seven cases to gather sufficient clues for the interviewer to conjecture a particular bias in the cast of mind (Spatial, Analytic or Harmonic) of the pupils. These have been listed but must be taken as tentative. Further work would be needed for them to be confident classifications. It may be that in view of the developmental stages of the pupils such casts of mind are difficult to detect. Item 22 was, nevertheless, retained on the checklist for the purpose of gaining further insights into this aspect of ability rather than for use in the identification process.

The categories that were developed for the items other than 18, 21 and 22 were not based on any preconception. They evolved from the clues that the pupils gave. The distinction between L, M and H was quite clear in all cases, in particular between M and H, where the clues that allowed a H (high) classification were very strong. The most difficult decisions came in the case of the moderately able pupils when the interviewer had to choose between L and M. He took as his key role that of giving adequate opportunities for the pupil to show his ability and then turn the observations into classifications with the question 'What has the pupil shown that gives an indication of the higher category?' in mind.

The question of which aptitudes are necessary and which are sufficient in the use of the checklist for identification in the classroom is one that requires consideration. It should be remembered in this that certain items group together to measure a particular aptitude that Krutetskii has isolated. Indeed it should be remarked that the evidence from the validity work is that these categories, as reflected in the items that were investigated, are relevant to the mathematical abilities of third form (13/14 year old) pupils. It would appear that able pupils score highly compared with those of lesser ability on a large number of the items so that the decision as to which are necessary and which are sufficient items may not have to be made. It should, however, be the topic of further research and until then only subjective judgements can be used.

The question also arises concerning the completeness of the checklist. Two points can be made. First, Krutetskii's large scale research was the basis of the checklist and Krutetskii's findings were confirmed for the kind of pupil in this study. Krutetskii's list can thus be used with confidence and it can be conjectured that his research was unlikely to miss the relevant aptitudes. Secondly, during all observations on the children it was possible to classify all of the clues they gave according to the checklist items, so that no new category was revealed.

Reliability

There were three separate exercises that were considered necessary to assess reliability of the conclusions from this work.

a) Inter-observer reliability

An experienced mathematics teacher was asked to act as a second observer with the same object in mind as the research officer - to record the clues to ability that occured during the lesson against the appropriate pupils' name. After the lesson the observers compared records and scores on the mathematical profile. Inevitably there were differences in what was recorded from the informal part of the lesson when the observers had responded to the needs of different pupils. There was, however, 100% agreement between the observers regarding the interpretation of the clues that were recorded, in particular where both observers had recorded the same piece of information.

It was seen as an essential component of this activity that both observers were making their interpretations with regard to the norms for that particular class and the stage of development within the topic.

A sample of the clues that arose for this work is given in Tables 10 and 11.

TABLE 10

CLASSROOM OBSERVATION RECORD SHEET (1 LESSON)

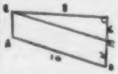
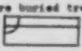
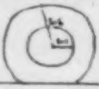
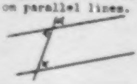
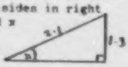
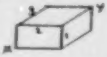
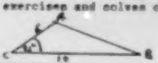
CHECKLIST CODE(S) AND LEVEL(S)	NATURE OF CHALLENGE	PUPIL'S NAME OR POSITION CODE	NATURE OF RESPONSE	CHECKLIST CODE(S) AND LEVEL(S)
	During first lesson on inverse proportion Qu. to class: 24 boys had 4 sweets each. If 32 wanted to share the sweets, how many each?		Immediately gives answer (most of class still thinking) $24 \times 4 \div 32$	1M 10L
	Developing an algorithm for inverse proportion by examples: 16 hrs. at 57 km per hour. What speed if same journey in 12 hrs? Teacher asks what do we mult. 16 by to get 12? Hence no. is $\frac{16}{12} \times 57$. No.hrs. Ave speed 16 <-> 57 12 <->		Pupil asks why $\frac{12}{16}$ not $\frac{16}{12}$ - not content with the explanation, persists in asking for an explanation as teacher does other examples and after class told to start their work from exercises.	15
	In further example WAVE FREQU. 371 <—> 809 1500 <—> Teacher says you can remember it by 'mult. the big no. by little no.'		Pupil sees a fault here and says, what if the big nos. are in the top line?	15, 14
	Boy's average step 60 cm. Father's 72 cm. Boy takes 840 steps for certain distance. How many steps would father take?		Answers immediately $840 \times \frac{60}{72}$; says that questions 'are all same', but takes a little time to explain proof.	1M, 5L, 8, 10?
	14 men do job in 48 days. Rains 11 days. how many more men are required to complete in same time?		Could not solve alone. Followed proof when explained but could not make any jumps of logic or construct any of the steps.	not 1M, not 5 not 8
	 Pupil succeeded already on area parallelogram que. Now find area ABCD.		Got confused over which numbers to use.	not 1, 2.
	New topic - solving equations. Class shown x-2=6 By 'add 2 to both sides'.		After several examples like this pupil gets some wrong and says he was taught to take 2 over and change its sign and can only do it that way.	not 10? not 14 not 15
	Pupil asked to explain what a reciprocal was.		Explains by reciprocal of $2 = \frac{1}{2}$ and when asked for more general explanation fails, but just repeats examples.	not 8
	Pupils just started to solve simple inequalities like $3m + 1 < -1 + m$ where $m \in Q$		Pupil gives appropriate answers with ease and explains clearly the relevance of set notation in these problems.	7b
	$7x \geqslant 7$		Pupil tries to subtract 6x from both sides.	not 14 not 6b
	$4x - 28 \geqslant 0$		Pupil completely foxed, though able to solve others with nos. other than '0'.	not 8 not 6b

TABLE 10 cont.)

CLASSROOM OBSERVATION RECORD SHEET (1 LESSON)

CHECKLIST CODE(S) AND LEVEL(S)	NATURE OF CHALLENGE	PUPIL'S NAME OR POSITION CODE	NATURE OF RESPONSE	CHECKLIST CODE(S) AND LEVEL(S)
	First lesson on intersection of loci, pupil asked if he was reminded about any other maths topic he had done.		No answer to question, but when prompted with 'what about sets', straight away explained that this was exactly the same because loci were sets of points.	7a
	Pupil had successfully worked with intersections of loci. Now attempting a problem where buried treasure was to be found.  Ext q.1		Pupil completely lost and needs very careful description, follows with difficulty.	not 1H not 8 not 5
	Pupil asked to describe $\{(x,y) : 1 \leqslant x \leqslant 8\} \cap \{(x,y) : -1 \leqslant y \leqslant 6\}$		Accurate description (first attempt at problems of this sort).	1, poss. 10? 7b?
	Pupils on first basic exercise on working out probabilities (such as (P| even no. on dice)). Pupil asked. "If A' means 'not A' what is $P(A')$"?		Immediately $P(A') = 1 - P(A)$.	7b. 8
	The first examples on mutually exclusive events are being attempted. Pupil can see when $P(A \text{ or } B) = P(A) + P(B)$		After solving several examples she cannot produce examples when $P(A \text{ or } B) \neq P(A) + P(B)$. Followed some examples and could still not produce one herself.	not 8 not 16
	What is 10^{-3} ?		No idea at start, but followed a long explanation through powers and roots. Could then produce answers to any similar question including 10^{-xy}	5. 8.
	One lesson only on expansion of brackets such as $(a + b)(c + d)$. One example only of $(x + 2)(x + 1)$, pupil asked to factorise $x^2 + 5x + 6$		Pupil succeeds. Was not required to solve this problem but persisted.	1c, 11b 20
	After lesson on factorising easy examples such as $ax + ay$, $6a + 12b$, pupils asked $6b^2(a + 2a^2) + 3b(a + 2a^2)$		Pupil 2 managed to get there but only after being shown much simpler cases.	5. not [illegible]

CLASSROOM OBSERVATION RECORD SHEET (RANDOM SAMPLE)

CHECKLIST CODE(S) AND LEVEL(S)	NATURE OF CHALLENGE	PUPIL'S NAME OR POSITION CODE	NATURE OF RESPONSE	CHECKLIST CODE(S) AND LEVEL(S)
4L 6bL	Class have just been shown how to solve $x + 6 = 12$ by adding -6 to both sides. Teacher asks class how to solve $-3x = 21$	2	Says multiply both sides by $-1/3$	4L 6bL
		6	Class have just been taught how to multiply with logs. Pupil will not get on with work until teacher explains how logs. work.	15M
	Teacher explains how to find area of a ring.  $\pi \times 3.6^2 - \pi \times 2.1^2$	10	Pupil says 'isn't it easier to factorise this as ' $\pi(3.6 - 2.1) \times (3.6 + 2.1)$'	13M perhaps H but not a 7bM long solution
	Class have just done some simple geometry on parallel lines. 	4	Pupil suddenly discovers that he can use this to prove that the angles of a triangle add to 180°	7bH possibly 22H
16M 6bM	Class have just learned to use sine and cosine to find sides in right Class asked to find x 	5	Immediately comes up with $\cos x = \frac{1.3}{2.1}$ and works sum out in head and uses tables to find angle.	16M 6bM 21H
4H 7bH	First lesson on 3 - geom. Class asked to find xy, distance between corners of a box. 	12	Pupil not only solves this but gives general formula $d = \sqrt{a^2 + b^2 + c^2}$	7bH probably 10H 4H 8H
6aL/M 6bL	After one lesson of sines and cosines class asked to get on with a simple exercise.	2	Pupil finishes routine ex. quickly and then finds a more interesting question in misc. exercises and solves completely 	7bM (Pythagoras 6bH 1H 4H 8H
6CL	Teacher gives class a 'puzzle'. Count no. of squares on chessboard.	1	Pupil solves but also gives a formula for the more general case of n x n board.	6CL 8H
	Teacher trying to prove pythagoras' theorem on board. Uses a circular argument.	4	Pupil corrects the mistake easily, points out where it went wrong and gives alternative solution	12H 13H Prob. 14H 11aH
		17	Pupil cannot see how to solve problem to do with length of arc of a rail of small 'O' guage train set when the radius is given. Thinks he needs to know what 'O' guage means.	Low on 2L

b) Clues that could be recorded from pupils workbooks

It was apparent, both from discussins with teachers and from the attempts that were made to find clues to ability by scrutinising workbooks of pupils, that very little on the broad spectrum of mathematical abilities could be reliably recorded. For this reason a more formal exercise was carried out with an experienced mathematics teacher who was also a tutor at the Department of Educational Studies.

A selection of exercises from the textbooks that were used by the pupils in the study were examined by the research officer and the tutor and were categorised according to the challenge that it could reliably be said were presented by the questions. Discussion took place before this categorisation to agree what stage, in the development of the topic, the pupils would have reached. The categorising was done separately and categories given by the research officer and the tutor were then compared.

Examples are given below. There was complete agreement on the categories and the main conclusion was that individual pupils may show particular clues that were not expected, but that on the whole a very narrow range of clues could be reliably recorded from work book scrutiny if the pupil was not present during the scrutiny to allow him to be questioned on what he had written.

Exercise 2 (Modern Mathematics For Schools)

By adding suitable numbers to each side, solve the following equations, the variables being on the set of integers. Give the solution in the form 'x = 11'.

1. $x+4=15$ 2. $x+7=19$ 3. $y+4=1$

2. $z+5=-2$ 5. $6+p=13$ 6. $21+q=75$

7. $t+(-2)=8$ 8. $x+(-7)=13$ 9. $x+(-5)=-9$

10. $2y=y+4$ 11. $5p=4p-10$ 12. $3m=8+2m$

13. $2x+3=x+7$ 14. $3t+1=2t+4$ 15. $2y+7=8+y$

16. $8v+10=7v+5$ 17. $3x-2=2x+5$ 18. $3w-2=2w-5$

19. $3p-44=6+2p$ 20. $5x+10=15+4x$ 21. $5y-4y+7=19$

22. $3x+12-2x=12$ 23. $5y+10-4y=-10$ 24. $4z+4=3z-9$

Checklist Code Levels All 6a(L) to 6a(M)

Exercise 3a (New General Mathematics)

Nos. 1-10 refer to $\triangle$ABC in which $\hat{B} = 90^\circ$. In each case find the length of the side which is not given.

1. AB = 6m, BC = 8m
2. AB = 8cm, BC = 15cm
3. AC = 26m, AB = 10m
4. AC = 25cm, BC = 24cm

16. A goat is tethered by a rope 13m long to a post which is 12m from a hedge. What length of hedge is within reach of the goat?

17. The heights of two vertical masts are 55m and 66m, and they stand 60m apart on the horizontal ground. Find the length of a taut wire (assumed to be straight) joining the tops of the masts.

Checklist Code Levels Q 1-4 6a(L)
Q 16/17 6b(L to M), 1 (L to M)

Exercise C (SMP Book E)

The first five questions are reprinted from Exercise A. Use the table of co-ordinates to answer them, and then check back on the accuracy of your scale drawings in Exercise A.

1. A plane is climbing at an angle of 20° to the horizontal. What is its increase in height when it has gone 1000 m?

2. A 20 m ladder leans against the wall of a house so that the angle between the ladder and the ground is 50°. How far is the foot of the ladder from the wall? How high is the top of the ladder above the ground?

3. The jib of a crane is 30 m long and makes an angle of 70° with the horizontal. How high is the top of the jib above the ground?

4. An escalator in a large store is 15 m long. It makes an angle of 30° with the horizontal. What is the distance between the two floor levels?

5. A ship sails 100 km on a bearing of 050°. (What angle does its course make with the east-west line?) How far north is it from its original position? How far east?

6. Figure 12 shows the cross-section of a roof of a house. What is the width of the house?

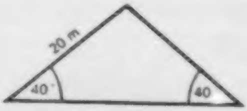


Fig. 12

7. A see-saw is 4 m long and inclined at 20° to the horizontal. How much higher is the upper end than the lower end?
If the lower end is 0.5 m above the ground, what is the height of the upper end above the ground?

Checklist Code Levels	Q 1-4	6a(M), 1(L)
	Q 5-7	6b(L to M), 1(L)

It was felt that workbooks were of limited use in the search for clues to ability so that any clues that it was necessary to follow up a large number of possible indications of ability with the individual pupils in class.

c) Reliability of identification of the top 10% and 5%

Even after our careful work connected with validation of the checklist items and reliability of the method we thought it necessary to look more closely at the pupils who were classified from the classroom observations to see if confirmation could be found for our conclusions.

A stratified random sample was taken from each of four of the classes who had been observed. In each of these classes two pupils were taken from the bottom of the group (according to checklist assessment), two from the middle, and two from the top 5%. These pupils were then given the series of mathematical challenges that has been described at the beginning of this appendix. Table 12 summarises the finding from this work. One important point regarding interpretation is that pupils in the same class could be given similar challenges, related to their common experiences, so that classifications are relative to norms for that class.

We can see that there is no conflict between what would be deduced about the randomly selected pupils from Table 12 and what was deduced from classroom observation, in the way the pupils were rank ordered or ability. The profiles are much the same as were reported for these pupils in Chapter Twelve. This, of course, is not really surprising, because during the active phase of observation the kinds of questions that were put to children in the classroom were modelled on those that were used for this reliability check.

One striking conclusion is that the separation of pupils into the percentile categories was not difficult. We might have expected to have needed to make value judgements about which scales were most relevant to high ability in order to separate pupils with different strengths and weaknesses across the profile. We found, however, that where differences existed they were in general in favour of one of the two pupils in a comparison on a considerable number of items. In the reliability check the only difficult decision concerned pupils 08104 and 08111 who were similarly ranked on a large number of items. It is further intersting to compare 08104 with, say, 107211, who was in another school. 08104 would have easily been in the top 10% of school 10, but relative to highly able contemporaries in his own school he was just below the 90th percentile on overall assessment.

All three reliability exercises gave results which strongly support the method of approach adopted and the interpretation of results that was used in the study based on classroom observation and reported in Chapter Twelve.

TABLE 12

CHECKLIST LEVELS FOR RELIABILITY EXERCISE

Identification level from observational work	Pupil code	Checklist item 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Below 90th Pctile	10102	M	L	L	M	M	M	L	L	L	L	L	M	M	L	M	L	M	M	L	-	-	-
Top 10%	10106	H	M	M	H	H	M	H	M	M	L	M	M	M	M	-	M	M	M	M	H	-	-
Top 5%	10121	H	H	H	M	H	H	H	M	H	M	H	M	M	H	H	H	H	H	H	H	-	-
Top 5%	10123	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	-	-
Below 90th Pctile	10131	M	L	L	L	L	L	L	L	L	L	L	L	L	L	M	L	L	L	L	L	-	-
Below 90th Pctile	10132	M	M	M	L	L	L	M	L	L	L	M	M	L	L	L	L	L	L	M	L	-	-
Below 90th Pctile	07102	M	M	M	L	M	M	M	L	M	L	M	L	L	M	L	M	L	L	M	L	-	-
Below 90th Pctile	07103	M	L	L	L	L	L	M	L	L	L	M	L	L	L	M	L	M	-	M	M	-	-
Top 5%	07120	H	H	H	M	H	H	M	H	M	M	H	H	M	M	H	M	H	H	H	M	-	-
Top 5%	07121	H	H	H	H	H	H	M	H	M	M	M	H	H	M	M	M	H	H	M	H	-	-
Below 90th Pctile	07124	L	L	L	L	M	L	L	L	L	L	L	L	L	L	L	L	L	L	M	-	-	-
Below 90th Pctile	07125	M	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	-	-
Below 90th Pctile	07209	M	L	L	M	L	L	L	M	L	L	L	L	L	L	L	L	M	M	M	M	-	-
Below 90th Pctile	07210	M	L	L	L	M	L	M	L	M	L	L	L	L	L	L	L	M	L	M	M	-	-
Top 10%	07211	H	M	M	H	H	H	M	H	H	M	M	H	M	H	M	M	M	M	M	H	-	-
Top 5%	07214	H	H	H	H	H	H	H	M	H	H	H	H	H	M	H	H	H	H	H	H	-	-
Top 5%	07224	H	H	H	M	H	H	M	H	H	M	M	M	M	M	M	H	H	H	M	M	-	-
Below 90th Pctile	07226	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	-	-
Top 5%	08101	H	H	H	H	H	H	H	H	H	M	H	M	M	H	H	H	H	H	H	M	-	-
Below 90th Pctile	08104	H	M	M	M	M	M	H	M	M	M	M	M	H	M	M	H	M	M	M	M	-	-
Below 90th Pctile	08107	M	M	H	H	M	M	L	L	M	L	L	L	L	M	M	L	M	M	M	M	-	-
Top 5%	08108	H	H	H	H	H	H	H	H	H	M	H	H	H	H	H	H	H	H	H	H	-	-
Top 10%	08111	H	M	M	H	H	H	H	M	M	H	H	M	M	M	M	H	M	M	H	H	-	-
Below 90th Pctile	08125	M	M	M	H	H	M	M	L	L	L	M	M	L	L	L	M	L	L	M	-	-	-

List of challenges from which appropriate choices were made during the testing sessions.

The checklist items that were normally monitored by means of the responses to these challenges are shown in brackets. The order of the challenges was not necessarily the order given here.

1. As a preliminary to the challenges general discussion of recent work covered from which the examiner noted down the topics. When a particular scheme such as S.M.P. had been followed to a particular book it was possible to get precise information quickly. If a particular pupil had received extra enrichment, closer questioning was needed.

2. In relation to work covered the pupil was asked to describe how to solve a particular sort of problem such as the solution of a simple equation, the solution of a quadratic, describe what a locus was, find the probability of an event. It was possible to see if the pupil was capable of describing the method generally or if he had to resort in the end to specific examples of the kind that he had practiced.

   (8, 17, 18)

3. The pupil was asked to say if a particular topic reminded him of others, if it was important to have knowledge of other topics before problems could be tackled. For example LOCI has links with graphs, co-ordinates, equations, sets, functions, PROBABILITY links with proportion, fractions, graphs, sets, INEQUATIONS with equations, sets, RATIO to fractions.

   (7)

4. Several new methods were demonstrated to the pupil which were not previously known but which they could tackle confidently because of their present expertise. For example:

   The solution of simultaneous equations.
   The solution of quadratic equations.
   The method of proof of the circle theorems.
   The method of induction.
   The derivation of the Cosine rule or sine rule.

The demonstration of $\sin^2 x + \cos^2 x = 1$

The demonstration of $\sin (A+B) = \sin A \cos B + \cos A \sin B$.

There were many opportunities here, depending on the pupil. The final aim was that he should achieve competance and confidence with at least one geometric and one algebraic investigation. For some pupils it would be possible to test flexibility in following a proof without harming their understanding. With others it was necessary to go very carefully so that each step was understood and repeated if necessary. These demonstrations were done early in the session and returned to as the final piece of work where the pupil was asked to describe the method in general terms and/or do some examples.

(4, 5, 6, 9, 10, 17,
possibly others such as
11, 14, 15, 16, 22).

5. If it was not possible in other work such as 4 above a test of flexibility was given where the examiner and the pupil together reasoned out a method and the examiner deliberately led the pupil along paths of logic that had to be retraced and deliberately made jumps of logic, small at first but of increasing difficulty, so that the pupils' flexibility of thought could be adequately challenged.

(5, 9, 10, 11, 12,
possibly 13, 14, 15, 16).

6. The following series of problems which tested the understanding of information make use of Krutetskii's own series of problems with appropriate modifications and additions. Each question was typed onto a separate card so that the pupil only saw and attempted one question at a time. The verbal instructions were along the following lines:

Understanding of information

These are presented in two series. First a guide to how the challenge was posed is given for each series and then the problems are listed.

Series 1 (Unstated question)

a) In each of these problems you are told the information but the question has been ommitted. You must say what you think the question should be.

b) (Abstract). What could we find out from the following information?

c) (Geometrical). What can we find out from the diagram and the information in each question.

Series 2 (Problems with redundant or superfluous information).

To solve a mathematics problem we need the right sort of information. In each of these questions there is either just enough information, not enough information or more information than necessary. You must say which sort of problem each of these is. You need not solve the problem.

(Demonstrate with 3 examples).

Series 1a

Example 1

A man earns £5 per hour. He works for 10 hours.

Example 2

A box contains some beads. John removes 12 beads. Jane removes 6 beads. Anne puts 4 beads into the box. There are then 11 beads in the box.

1. An athlete takes 2 minutes to run round a race track. He runs for 6 minutes.
2. A train has 12 coaches. Each coach can hold 40 passengers.
3. A bus is carrying 25 passengers. It picks up some more passengers at one stop and 14 get off at another. It is then carrying 32 passengers.
4. A bath can hold 80 gallons of water. A tap is turned on and lets in water at 4 gallons per minute, but the plug has been removed so the water escapes at 2½ gallons per minute.
5. In a bag are 220 beads. The beads are in 3 sizes. There are 10 more small and medium beads together than large beads and 5 more small beads than medium beads.
6. 40 weights of 2 kg or 4 kg each are in a box. The total of all the weights in the box is 96 kg.
7. Bill and Ben together have £500. When Bill gives Ben £25 they each have £250.
8. The number of pound notes that I would need to pay a bill is 3 less than the number of 50p pieces that I would need to pay the same bill.
9. A boy has as many sisters as brothers, and one of his sisters has half as many sisters as brothers.

Series 1a cont.

10. A man paid £5 for some records. If the records had been 10p cheaper he could have bought 4 more records.

11. Three years ago Peter was three times as old as Paul. In twelve years Paul will be half of Peter's age.

12. Two pies cost half as much as six cakes. Two cakes cost the same as four buns.

13. A car starts with a full petrol tank which holds 14 gallons. At a certain point on the journey the driver observes that he has remaining ¼ of what he has already used.

14. A car travels between two towns at 45 mph. A bicycle goes on the same journey at 12 mph. The car leaves 2 hours after the bicycle but arrives 3 hours before the bicycle.

15. After a party, a quarter of the children did not want a bar of chocolate and a third did not want an apple. Half of them took both a bar of chocolate and an apple while two children took neither chocolate nor apple. One apple and one bar of chocolate had been bought for each child.

16. A boy runs up a moving staircase that is going downwards. The boy runs up four steps per second. A man begins to walk down another staircase with the same number of steps and which is not moving. He walks down 3 steps per second and starts from the top 4 seconds after the boy started from the bottom. The man reaches the bottom after 20 seconds. The boy gets to the top at the same time.

17. 20 people wanted to be driven from a hotel to an airport. Two taxis took three people each. Ten people went by bus and the rest walked. Taxi drivers charged £1.50 for each journey. The 20 people decided to share the total amount spent on travel and each paid 75p.

18. Ian works for four hours each day and works on six days per week. Brian earns £120 per week and is paid £1 less than Ian for each hour that he works. Ian works six hours more every week than Brian and Brian works on only five days per week.

19. Five numbers are written down in a row. The sum of the first three numbers is 60. The sum of the last three numbers is 90. The five numbers add up to 110.

20. A box contains 50 pencils whose colours are either red or white or blue. Twice the number of red pencils added to three times the number of white pencils is 142 less twice the number of blue pencils.

Series 1b.

1. The price of a table and chair. The price of a chair.

2. The number of litres of petrol used by a car on a journey. The number of miles that the car travels per litre.

3. The number of revolutions made by a wheel when it rolls from A to B. The distance between A and B.

4. The number of days that a woman lived. The date that she died.

5. A man's weight in kilogrammes. The number of pounds that make a kilogramme and the number of pounds that make a stone. How much more the man's weight is than his son's in stones and pounds.

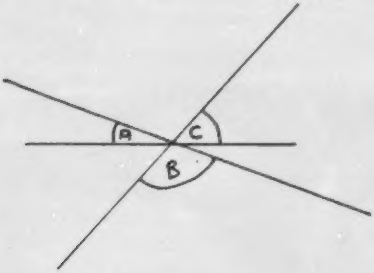
6. A has x pounds more than B. B has y pounds less than C. Y is Z more than X.

7. Baked bean tins are packed in boxes of two sizes. All tins are the same weight. The total weight of boxes on a lorry is known. The weight of a large box of beans is known. The number of tins in a small box is known. The number of large boxes on the lorry is known and the total number of boxes is known.

8. The distance between two trains that are travelling in the same direction at steady speeds. The further distance that the following train goes before it overtakes the other. The time that the second train takes to overtake the first.

Series 1c.

1.



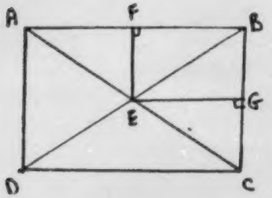
$\hat{C}$ is $^3/_5$ of a right angle.

$\hat{A}$ is half of $\hat{C}$

Series 1c. cont.

2. One angle of a triangle is 30° larger than another and the third is 20° smaller than the first.

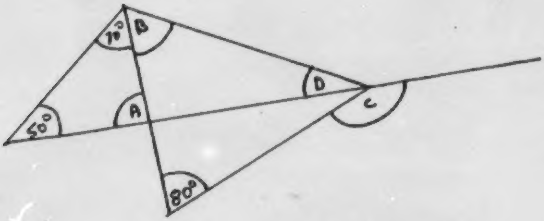
3.



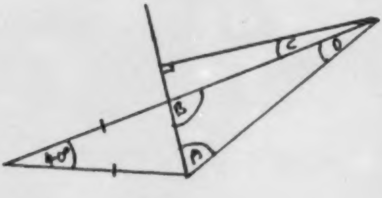
The perimeter of rectangle
ABCD is 80 cm.
EG is 12 cm. more than EF.

4. Which angles could be calculated in the following figures?

a)



b)



Series 2a.

Example 1

A man bought 10 stamps and received 40p change. How much did each stamp cost?

Example 2

A ten year old girl ran for five minutes at 10 mph. How far did she run?

Series 2a cont.

Example 3

Two people each pay half of a bill that comes to £10. How much does each pay?

1. Some children go into a sweet shop and each buys a bar of chocolate costing 12p. What is the total cost?

2. In a box there are some beads. There are 10 more red beads than blue beads and 8 more blue beads than green beads. How many of each kind are there?

3. Two men count how much money they have in their pockets. One is two years older than the other and has four more pounds. Together they have twenty five pounds. How much money does each man have?

4. An electric kettle takes ten minutes to boil two pints of water. How long would it take to boil six cups of water in the kettle?

5. A class of schoolchildren were given pens and pencils - 80 altogether. Pens cost 12p and pencils cost 9p. How many of each kind were given out.

6. Two chairs and a table together cost £50. Three chairs cost £40. How much does a table cost?

7. Seven men dig a hole in seven minutes. Six men dig a hole that is exactly the same in six minutes. How long would it take five men to dig the same hole?

8. 40 vehicles are in a car park - motorcycles and cars. Altogether there are 100 wheels and 80 number plates. How many of each kind of vehicle are there?

9. A woman has £13 in pound notes, 50p and 10p coins. She has 4 more pound notes than 50ps and 3 times as many 10ps as 50ps. How many of each sort of money does she have?

10. A boy had a few pence. When he was given 14p more he bought 4 pencils, spending all his money and paying twice as much for each as he had to begin with. Before he was given the extra money he could not afford even one pencil. How much did each pencil cost?

11. Four weights together weigh 40 kg. Find the weight of the heaviest one if each one is three times as heavy as the next and the lightest one is twelve times less than the two middle ones together.

12. A library contains 3250 books on Mathematics, Physics and Chemistry. There are 20% more Mathematics books than Physics books. How many of each kind are there?

13. Half a piece of string and ½m. were cut off. Then half of the remainder and ½m. were cut off what was left. Finally half of what was then left and ½m. were cut off. How long was the string at the start?

Series 2a cont.

14. Two cities are 200 miles apart. A train leaves the first city and travels at a steady 40 mph. Half an hour later a second train departs from the same city and travels at a steady 60 mph. The second train is twice as long as the first. How far from the first city will the second train be overtaken?

15. 3 cups and 1 saucer cost the same as 2 plates. 1 cup and 3 saucers cost the same as 4 plates. How many saucers cost the same as 15 cups?

16. A library needs to repair some English, French and German books whose numbers are in the ratio 3:2:1. One firm could do the job in 20 days, a second in 30 days, a third in 60 days. How long would it take if the firms worked simultaneously?

17. A dog is chasing a fox which is 30 m. away. A dog's stride is 2 m. while a fox's is only 1 m. In how many strides will the dog catch the fox?

18. An empty jar is three times as heavy as an empty bottle. When the jar is filled with water it weighs 500 g. How much will the bottle weigh when it is filled with water?

19. Two signal boxes are 4 miles apart. A train going at 40 mph passes one at the same instant that another train going at 60 mph passes the other. How long before the engines pass one another?

20. In a cafe, milk costs twice as much as tea and coffee costs 10p more than milk. A tea cup holds ¼ as much as a coffee cup. Four men each choose one drink and the bill comes to £1. How many coffees were bought?

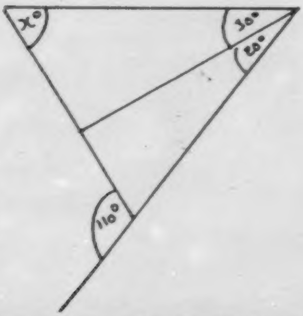
Series 2b.

1. Calculate the side of a square of area 64 cm^2. Calculate the side of a rectangle of area 36 cm^2.

2. In an isosceles triangle the lateral side is less than the base. Its perimeter is 31 cm. What is the length of the triangle's sides?

3. Given 2 circles. The radius of one is 3 cm. and the distances between their centres is 1[illegible] cm. Do the circles intersect?

4. Find x°.



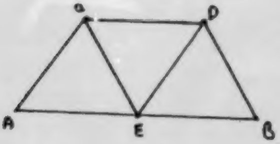
Series 2b cont.

5. Given an isosceles triangle with one side 2 cm. and another 10 cm., the third side equal to one of these two. Find the third side.

6. Two sides of an isosceles triangle are in the ratio 3:8. Find the sides if the perimeter is 38 cm. and one side is 10 cm. longer than the other, giving the sides as whole numbers.

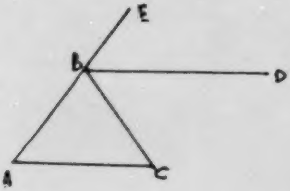
7. A straight line is drawn through the vertex of an angle and outside of it forming an angle $^{d}/3$ with one of its sides. Determine the size of the angle formed by the line and the other side of the given angle.

8. Given three congruent equilateral triangles with sides of 6 cm., show AEB is a straight line.

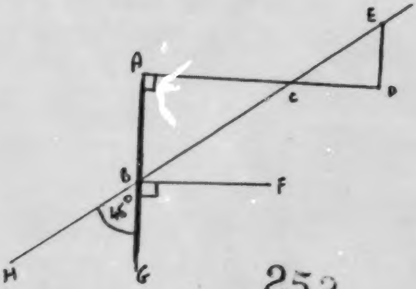


9. The sides of a triangle are in the ratio 5:4:3. Find the lengths of its sides.

10. Given AB = AC = BC. △ ABC is equiangular. BD is parallel to AC. Show BD bisects <EBC.

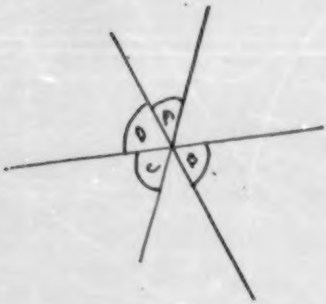


11. Show AB = AC



Series 2b cont.

12.



$\hat{A} = {}^2/_3\,\hat{B}$
$\hat{C} = 42^\circ$
Find $\hat{D}$.

(1, 2, 3)

7. The following illogical arguments were demonstrated step by step and the reaction to the invalid logic was observed. If nothing was said during the demonstration the pupil was given the opportunity to reflect on the result and to give any comments. It is worth remarking that one unexpected observation from this was that a pupil's aptitude concerning persistence with a problem could be monitored. The pupil might accept it as a mathematical challenge to find the place where the logic went wrong.

(14, possibly
13, 15, 20).

a) An illogical conclusion from inequalities.

Suppose we wish to solve $-3x + 2 < 11$
Then subtracting 2 from both sides $-3x < 9$
Dividing by - 3 $x < -3$

b) Demonstration that 45 = 0 (From Northrop E.N., Riddles In Mathematics,
Pelican 1944)

First remind how to subtract with 'borrowing'.

e.g. $\begin{array}{r} 461 \\ \underline{137} \\ 324 \end{array}$

Explain how 'borrowing' works.

Now consider $\quad 9+8+7+6+5+4+3+2+1 = 45$

$\qquad\qquad\qquad \underline{1+2+3+4+5+6+7+8+9} = \underline{45}$

Subtract, following
'borrowing' on left

$\qquad\qquad\qquad 8+6+4+1+9+7+5+3+2 = 0$

Thus 45 = 0

c) Proof that each person had about 1,000,000 ancestors 600 years ago and consequently there were 1,000,000 times as many people on earth then.

Each of us has 2 parents.

Each of them had 2 parents and so on, so that

1 generation ago we had 2 ancestors

2 " " " " 4 "

3 " " " " 8 "

If we go on doubling, 20 generations ago we each had 1,040, 000 ancestors (2^{20}).

A generation is about 30 years. 20 x 30 = 600 so

600 years ago we each had over 1,000,000 ancestors. (Northrop E.N. op. cit.).

d) Same problem, 2 methods, different answer (Frohlichstein J. Mathematical Fun, Games and Puzzles, Dover 1967).

When we multiply:

$$\begin{array}{r} 16.5 \\ \underline{\text{X}12.5} \\ 825 \\ 330 \\ \underline{165} \end{array}$$

our answer is: 206.25

However, when we multiply

$$\begin{array}{rl} 16\tfrac{1}{2} & \\ \underline{\text{x}12\tfrac{1}{2}} & \\ 8\,{}^{1}/4 & (1/2\text{x}16\tfrac{1}{2}) \\ 6\,{}^{1}/4 & (1/2\text{x}12\tfrac{1}{2}) \\ 32 & (16\text{x}2) \\ \underline{16} & (16\text{x}1) \\ 206\tfrac{1}{2} & \end{array}$$

we get 206½ or 206.5, which is not the same as the decimal answer. What has happened to the other 1/4?

e) To show any two unequal numbers are equal. (Northrop op.cit.)

Let a and b be two unequal numbers, and let c be their arith metic mean, or average (for example, if $a = 2$ and $b = 4$, then $c = (a+b)/2 = 3$). Then

$$\frac{a+b}{2} = c, \text{ or } a + b = 2c. \qquad (1)$$

Multiply both sides by $a - b$:

$$a^2 - b^2 = 2ac - 2bc. \qquad (2)$$

Add $b^2 - 2ac + c^2$ to both sides:

$$a^2 - 2ac + c^2 = b^2 - 2bc + c^2 \qquad (3)$$

Both sides of (3) are now perfect squares and can be written in the form

$$(a - c)^2 = (b - c)^2 \qquad (4)$$

Take the square root of both sides. Then

$$a - c = b - c, \qquad (5)$$

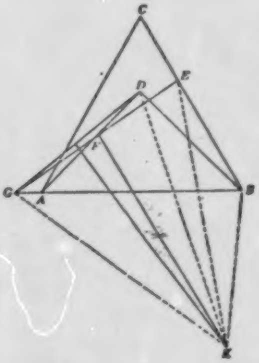
or

$$a = b.$$

We started with the assumption that a was not equal to b and have come to the conclusion that a is equal to b.

f) To prove $45^\circ = 60^\circ$, or $3 = 4$ (Northrop op. cit.)

On side AB of equilateral triangle ABC as hypotenuse construct an isosceles right triangle ABD. We shall prove that ∠ABC, which is 60°, is equal to ∠ABD, which is 45°.



On BC lay off BE equal to BD. Mark F, the mid-point of AD, and through E and F draw a line which intersects BA produced in G. Draw GD. Then construct the perpendicular bisectors of GE and GD. Since GE and GD are not parallel, the perpendicular bisectors must meet at some point, say K. Connect K with G, D, E and B. Our main job now is to show that triangles KDB and KEB are congruent. Note first that KG = KD and KG = KE. (Any point in the perpendicular bisector of a line is equidistant from the ends of the line.) Hence KD = KE. Furthermore, BD = BE by construction, and BK is a common side. Therefore triangle KDB is congruent to triangle KEB, whence ∠KBD = ∠KBE. If now we subtract the common portion KBA from each of these angles, we must conclude that ∠ABD = ∠ABC, or that 45° = 60°, or that 3 = 4.

8. Selections were made from the following series as appropriate for the pupil in order to see how readily the pupil could give ideas as to how to proceed with a solution. This was an open ended challenge so that the interaction with the pupil could lead to a variety of different clues, though the pupil clearly understood that there was no necessity to give a complete solution, but a few ideas to get started was all that was necessary.

(4, usually 1, 2 or 3,
12, 7).

Series 3

1. A cyclist, who was a shade under six feet tall, started out from a place, which for lack of any knowledge of its name we will call 'A', at half-past ten one sunny morning. He cycled due south for a distance of four miles; then, turning west, he cycled a further two miles. Turning south once more he continued his journey for another eight miles and then making a complete turn to the left, he sped along due east at an average speed of eleven miles an hour. Reaching 'B' exactly one hour after turning left, he dismounted and enjoyed a leisurely lunch at the Rose and Crown. After refreshing himself he enquired of the landlord the quickest route back to 'A' and was surprised to learn that there was a path through the woods which ran, straight as a crow would fly, from 'B' to 'A'. He followed this path home and arrived there a little after 3 o'clock. What was the distance he cycled after lunch?

(J.E. Thompson 'Puzzle Playtime'
Wheaton & Co. 1961)

2. For what values of 'x' and 'y' is the following equation true?

$$2x(1 + y) = x^2 (1 - y).$$

(J.E. Thompson op.cit.)

3. John was passing the time which he should have devoted to the more serious business of study by playing with figures. He wrote down a number, subtracted 1 from it and doubled his answer. From the new number he thus obtained he subtracted 2 and once again doubled his answer. Subtracting 3 from this last product he multiplied the result by 3 and, to his surprise found that his final result was nine times as great as the number he had first written down. What was this first number?

(J.E. Thompson op. cit.)

4. Last Monday morning while Mr. Jones, our village grocer-cum-baker-cum-postmaster-cum-nearly-everything-else was rearranging his counters his automatic scales fell over and were damaged. Trade had to go on so Mr. Jones turned out his old scales but he could find only three weights. If these enabled him to weigh any whole number of pounds from one to thirteen what were their values?

(J.E. Thompson op. cit.)

Series 3 cont.

5. One very sleepy afternoon last summer two cars were approaching each other on a straight stretch of a well known motorway. Each car was travelling at the surprisingly low speed of thirty miles an hour. When they were sixty miles apart a bee that had been sleeping on the bonnet of one of the cars suddenly awoke and flew off towards the other car at a speed of seventy miles an hour. On reaching the second car it immediately turned round and sped back still at seventy miles per hour to the first car. Turning round again it flew back to the second car. If the bee kept up this somewhat energetic performance until the cars met and all the while maintained its steady seventy miles an hour, how far did it fly?

(J.E. Thompson op. cit.)

6.

$$\begin{array}{r} x8 \\ \underline{xx} \\ xxx \\ \underline{xx40} \\ xx36 \end{array}$$

In the accompanying multiplication sum the figures indicated by 'x' were indistinct in the copy and hence it was impossible to reproduce them here. Can you find them?

(J.E. Thompson op. cit.)

7. Find a number made up of three digits which become inverted when added to 198 and which when added individually together give a sum equal to 21 if the sum of the first and second digits is greater than the third by 3.

(J.E. Thompson op. cit.)

8. In a village 10% of the inhabitants had lost one leg apiece. Of the remainder, one half were the proud possessors of shoes which, although they were not in the best state of repair, their owners wore on all possible occasions. On the first day of each month the entire village assembled before the Head Man's hut and at these assemblies the one-legged natives were each given a shoe to wear. During one of these ceremonies it was noticed that just 100 shoes in all were being worn. How many people lived in the village?

(J.E. Thompson op. cit.)

9. On the death of old Mr. Robinson it came as a matter of some surprise to a number of his cronies that he had so much property. Quite naturally he left a considerable part of it to his three sons and, being a precise man in life, he left precise instructions concerning the distribution of this part in particular. Tom, his eldest son, was to enjoy three-fifths of it and John was to receive twice as much as his brother Harry. If, on selling out that part of the old man's property that was to form the sons' inheritance, it realised £30,000 what did Harry's share amount to?

(J.E. Thompson op. cit.)

Series 3 cont.

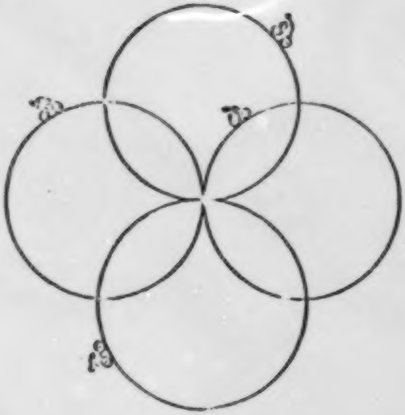
10. Seven applewomen, possessing respectively 20, 40, 60, 80, 100, 120 and 140 apples, went to market and sold all their apples at the same price, and each received the same sum of money. What was the price?

(Dudeney, H.E. Puzzles and Curious Problems Fontana 1967)

11. A train is travelling at the rate of sixty miles per hour. A passenger at the back of the train wishes to walk to the front along the corridor and in doing so walks at the rate of three miles per hour. At what rate is the man travelling over the permanent way? We will not involve ourselves here in quibbles and difficulties similar to Zeno's paradox of the arrow and Einstein's theory of relativity, but deal with the matter in the simple sense of motion in reference to the permanent way.

(Dudeney, H.E. op. cit.)

12.



The four circles represent cinder paths. The four cyclists started at noon. Each person rode round a different circle, one at the rate of six miles an hour, another at the rate of nine miles an hour, another at the rate of twelve miles an hour and the fourth at the rate of fifteen miles an hour. They agreed to ride until all met at the centre, from which they started, for the fourth time. The distance round each circle was exactly one-third of a mile. When did they finish their ride?

(Dudeney, H.E. op. cit.)

Series 3 cont.

13. Using the French Tricolor as a model, how many flags are possible with five available colors if two adjacent stripes must not be coloured the same?

(Dunn, A. Mathematical
Bafflers, Dover 1980)

14. A cubic box with 1-foot edges is placed flat against a wall. A ladder 15 feet long is placed in such a way that it touches the wall as well as the free horizontal edge of the box. Find at what height the ladder touches the wall.

(Dunn, A. op. cit.)

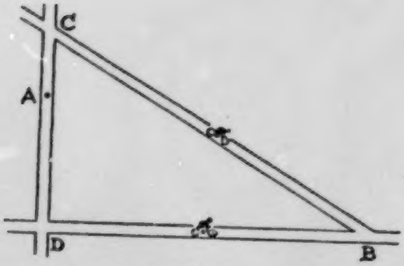
15. A certain physicist, who is always in a hurry, walks up an upgoing escalator at the rate of one step per second. Twenty steps bring him to the top. Next day he goes up at two steps per second, reaching the top in 32 steps. How many steps are there in the escalator?

(Dunn, A. op. cit.)

16. In an examination 60 candidates offered Maths, 80 offered English and 50 offered Physics. If 20 offered Maths and English, 15 English and Physics, 25 Maths and Physics and 10 offered all three how many candidates entered the examinations?

(Greer, A Complete O level
Mathematics, S. Thornes 1973)

17.



Two army motorcyclists, on the road at Adjbkmlprzll, wish to go to Brczrtwxy, which for the sake of brevity, are marked in the accompanying map as A and B. No, Pipiposs said: "I shall go to D, which is six miles, and then take the straight road to B, another fifteen miles." But Sliponsky thought he would try the upper road by way of C. Curiously enough, they found on reference to their cyclometers that the distance either way was exactly the same. This being so, they ought to have been able easily to answer the General's simple question, "How far is it from A to C?"

(Dudeney, H.E. op. cit.)

9. When adequate clues had been gathered from 8, one or two problems were found to interest the pupil and attempts were made to go further with the solution, encouraged by the examiner. This was usually before a break, so that the pupil could be encouraged to persist in seeking a solution during the break.

(20)

10. It was possible in some of the other work to observe checklist item 7 where the pupil's ability to use knowledge of maths topics of one sort transferred to problems in another maths area. Additional challenges were availabl if clues had not been found elsewhere such as:

a) Proof that $\sin^2 x + \cos^2 x = 1$ (Pythagoras)

b) Finding the equation of a circle (Pythagoras)

c) Addition of vectors (Sine rule, Cosine rule).

(7)

11. Krutetskii's problems to check the aptitude of direct to reverse train of thought were used almost directly. The pupil was asked to think aloud and the time from understanding the problem to solving it was noted. The problems were typed on cards with the second problem on the opposite side to the first so that it was possible to give the second problem immediately after the first.

(16).

Series 5

1. Sixteen litres of water were poured into a tank, filling the tank to two-fifths of its volume. What is the volume of the tank?

   Water was poured into a tank of 80-litre capacity to two-fifths of its volume. How many litres of water were poured into the tank?

2. A saw in a sawmill saws off a 1 m piece of log every minute. How many minutes will it take to saw up a 16 m log?

   In 3 minutes a log is sawed up into half-metre pieces, with each cutting taking 1 minute. Find the length of the log.

Series 5 cont.

3. A factory with a plan for an output of 1,280 machine tools exceeded the plan by 12.5%. How many tools did it turn out?

   A factory that exceeded its quota by 2% turned out 51 tools. What was the planned output?

4. Using 740 sheets of paper, 140 thick and 100 thin notebooks were made. How much paper was used for thick and thin notebooks separately if 1 more sheet of paper went into a thick notebook than into a thin one?

   Using 620 sheets of paper, 200 notebooks were made. The notebooks were made of 5½ and 1½ sheets each. How many notebooks of each kind were made?

5. A mother is 3 times as old as her daughter, 10 years from now she will be only twice as old. What is the mother's age?

   A father is 35 years old, and his son is 5. In how many years will the father be 3 times as old as the son?

6. How many days should a worker spend working to earn b pounds if he gets c pounds per day?

   How much does a worker earn in d days if he earns a pounds in one day?

7. The distance between cities A and B is x km. Two trains departed travelling towards each other. One went at a speed of a km per hour, the other at b km per hour. How long did it take before they met?

   The distance between two points is y km. Two trains departed from them, going towards each other, and they met in n hours. One train travelled at a speed of 40 km per hour. What was the speed of the second train?

12. Krutetskii's investigation of a pupil's ability to generalise from

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$\text{to} \quad (C + D + E)\ (E + C + D)$$

was available for appropriate pupils and the same system of giving progressively stronger hints of the type:

1. $(a + b)^2 =$
2. $(1 + \frac{1}{2}a^3b^2)^2 =$
3. $(-5x + 0.6xy^2)^2 =$
4. $(3x - 6y)^2 =$
5. $(m + x + b)^2 =$
6. $(4x + y^3 - a)^2 =$
7. $51^2 =$
8. $(C + D + E)(E + C + D) =$

was adopted.

A map of the route from $(a + b)^2$ to $(C + D + E)(E + C + D)$ was made. A route such as $2 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 8$ would indicate a moderate ability while $1 \rightarrow 8$ would indicate high ability. In cases where $(a + b)^2$ was a familiar expansion, then similar work with $\sin(a + b)$ and $\cos(a + b)$ to $\sin(C + D + E)$ was an alternative.

(8, 10)

13. Krutetskii's problem given below was available for gaining insights into generalization and curtailment of reasoning.

(8, 10)

> The pupils were to deduce, on their own, the rule for the short calculation of the square of two-place numbers ending in 5. The following directions were set for the pupils, and examples were given: $15^2 =$, $25^2 =$; $35^2 =$; $45^2 =$; $55^2 =$; $65^2 =$, $75^2 =$, $85^2 =$; $95^2 =$ In doing the appropriate calculations, the pupils were to attempt to construct an elementary algorithm for rapid and easy squaring of such numbers. This is not hard to do if all such numbers are conceived in a general form: $(10a + 5)^2 = 100a^2 + 100a + 25$. In which example did the pupil detect the guiding principle? By what means did he arrive at it?

Other of Krutetskii's series were tried with different pupils in order to attempt to get further insights (for example Series IX: Problems on Proof were given to one able child). In general these extra problems confirmed what had been found in other parts of the session for that child. Space does not permit the listing of all of these additional challenges.

14. Questions that could be tackled in more than one way were given to the pupils (a selection of Krutetskii's questions). (12)

Series 4

1. In how many ways can 78 pounds be paid if the money is in 1 pound and 5 pound notes?
2. In 4 classes there is a total of 118 pupils, including 70 in grades one and two together, 65 in grades one and three together, and 59 in grades two and three together. How many pupils are in grade four?
3. Find the sum of all the integers from 1 to 50.
4. Sailing with the current, a steamer does 20 km per hour; against the current it sails at 15 km per hour. To travel from A to B, it uses 5 fewer hours than when travelling in the opposite direction. What is the distance between A and B?

Series 4 cont.

5. Four litres of water at room temperature (15 C) were added to 3 litres of water at a temperature of 36°. What temperature was established in the vessel?

6. $113^2 - 112^2 =$

15. It was possible both at the end and at other times during the work to assess the levels of fatigue due to intense work. This was by several methods. Often the pupil would show loss of concentration and confirm that a short break would be beneficial. Pupils were quite honest about their levels of fatigue. Beneficial effects of a short break were observable to confirm this.

(19)

CHECKLIST OF MATHEMATICAL APTITUDES

No. ITEM

UNDERSTANDING OF INFORMATION

1. Able to understand the formal structure of a problem, seeing the relationship between the items of information even when there is a lot to assimilate.

2. Able to see when there's redundant information in a problem even when there is a lot to assimilate.

3. Able to see when there is not enough information in a problem and to see what information is needed for its solution even when there is a lot to assimilate.

PROCESSING OF INFORMATION

4. Able to suggest plausible ideas that might lead to the solution of a new sort of problem.

5. Able to follow the logic in a long mathematical proof.

6. Able to construct mathematical proofs:
   a) similar to ones already demonstrated;
   b) requiring different logic from that already demonstrated but on a familiar topic;
   c) on problems related to a new or unfamiliar topic.

7. Quick to see when a method of solution that has been applied to problems in one topic will solve problems of a similar nature in another:
   a) with help;
   b) independently.

8. Quick to see when examples of a mathematical result have a wider and more general application: focuses readily on these generalisations rather than on the particular examples.

9. Able to follow a logical explanation or proof when valid logical jumps are made that miss out or assume some of the intermediate steps of logic.

10. Displays an ability to make valid logical jumps when using mathematical reasoning accompanied by the ability to provide the intermediate steps.

11. When a method fails in a process of reasoning is able to try an alternative approach incorporating what has already been validly reasoned into the new structure:
    a) in following someone else's logic;
    b) in constructing a system of logic independently.

No.	ITEM
	PROCESSING OF INFORMATION cont.
12.	Is able, when possible, to give alternative methods of solving a problem.
13.	Is critical of inelegant solutions and strives in his/her own work for the most elegant.
14.	Is quick to notice an incorrect step of logic or an invalid conclusion.
15.	Is not content unless he/she has complete understanding of a mathematical concept, method or process of logic.
16.	Adapts readily from a direct to a reverse train of thought.
	MATHEMATICAL MEMORY
17.	Has a good memory for a method of solution. Remembers a proof by its logic rather than by rote learning.
18.	Has a good long-term memory for techniques developed for a particular topic. Requires little or no revision to regain previous levels of competance.
	GENERAL
19.	Shows little mental fatigue even after intense concentration on difficult problems.
20.	Is not content to leave an interesting problem unsolved, but keeps returning to it at odd moments until success is achieved.
21.	Can perform any arithmetical computation accurately and with ease.
22.	When attempting to solve new sorts of problems displays a strong tendency to interpret and solve the problem in: a) a spatial/geometric way; or b) an analytic/logical way.

Case studies of two testing sessions, one of a moderately able pupil, one of a pupil of known high ability

It was felt that the most useful way of reporting the two case studies was to consider each checklist item in turn and report the evidence that led to the classifications (L low, M medium or H high) attributed to these pupils. These pupils have been randomly selected from each of the groups.

1. Moderately able pupil MM9

ITEM 1 (M) Series 1a answered correctly 1-11, 13
incorrectly 12, 14-20

Series 1b answered correctly 1, 2, 4, 5
incorrectly 3, 6, 7, 8

Series 1c answered correctly 1, 2, 3
incorrectly 4

Slow to understand series 3 question 1, but after several attempts able to commence solving.

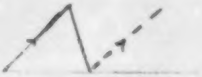
ITEM 2 (M) Series 2a correct 1-4, 6, 9, 15, 18
Incorrect 5, 7, 8, 10- 20.

ITEM 3 (M)

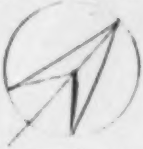
Able to sort out, after several attempts, the relevant information in series 3 problems.

ITEM 4 (M) Attempted to give an answer every time he was asked for an idea but not always a practicable one. The following (unfamiliar) proofs were worked with the examiner.

P1 Proof that the sum of the angles of a triangle add up to 180°. He suggested drawing and measuring but could go no further. When the examiner demonstrated the construction of drawing a parallel line and asked for the next step, he noticed some equal angles but could not complete the proof.



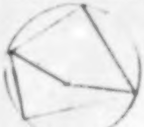
P2 Four of the five circle theorems were demonstrated by the examiner: angle at centre twice angle at circumference - suggested measuring. Only real idea during proof came when isosceles triangles pointed out.



Angles at circumference equal. Wanted to measure them again. Suddenly saw relevance of previous theorem after the hints 'hasn't the last thing we proved got something to do with this' and then the angle at the centre was drawn.

Angle in a semi-circle is a right angle. Saw this quite quickly.

Opposite angles supplementary. Had to be shown construction and given strong hints before he made use of first theorem



P3 How to solve simultaneous equations.

1) $3x + 2y = 7$
2) $2x - y = 6$

No real idea how to start. Kept rearranging one equation at a time, often with errors. Suggested last step of substituting the value of y back in equation 1 to find x

P4 Quadratic equations. Several attempts to start

$$x^2 + 2x + 3 = 0$$

such as dividing by x and arriving illogically at $x + 2 + 3 = 0$. No real ideas at any stage of the proof.

Ideas given for series 3 problems were:

Q1 Work on squared paper. Draw a diagram. (Plausible but couldn't put the ideas into practice without help)

Q2 Tried to multiply out and rearrange. No success. Did not spot trial and error.

Q4 Tried number combinations for adding only.

Q5 Could only suggest a step by step approach. Realised it was not a good approach. Tried to think of other methods. Suggested drawing graphs but could not follow it through.

Q6 Trial and error in a logical way getting easy numbers first. Succeeded.

ITEM 5 (M) Followed P1 to P3 with no difficulty. The proof for quadratic equations in the general form was followed only after several attempts and particular examples. Main difficulty with completing the square even on particular examples, final steps of rearranging and finding final form of square root. Difficulty associated with keeping the overall strategy in mind. Not really confident of overall plan even after following and understanding each step.

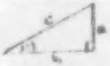
ITEM 6 (M) Could describe the sequence of logical steps to solve problems and prove results connected with his recent work such as inequalities and probability. Maths teacher confirmed ability within recent topic work. Had difficulty with all but one of the series 3 puzzles attempted. Gave no more than weak ideas, though clearly enjoyed the challenge of puzzles.

ITEM 7 (M) When asked 'what other topics were useful?' and 'what other topics do you think of?' when recent familiar topics were discussed some links were established independently and others after hints.

TOPIC	INDEPENDENTLY	AFTER HINTS
Inequations	Algebra, equations	Sets, co-ordinates, graphs
Ratio	Arithmetic, division	Fractions, ordered pairs, graphs
Probability	Fractions, Ratios	Sets, ordered pairs
Locus	Graphs	Equations, Sets, Co-ordinates

The pupil had previously shown an understanding of Pythagoras' theorem. The examiner asked if the pupil could prove $\sin^2 x + \cos^2 x = 1$. It was necessary to explain that $\sin^2 x$ meant $(\sin x)^2$. The pupil had no idea. The examiner drew a triangle.

The pupil was asked if it reminded him of anything. The pupil said trigonometry. The examiner labelled the triangle and wrote down $\sin x = \frac{a}{c}$ $\cos x =$



The pupil was asked if anything else came to mind looking at the diagram. The pupil did not know. The examiner wrote down $\sin^2 x + \cos^2 x = \frac{a2}{c2} + \frac{b2}{c2}$

but it was not until he transformed the right hand side to $\frac{a^2 + b^2}{c^2}$ and was asked 'what does $a^2 + b^2$ equal?' that the pupil remembered Pythagoras' theorem.

The pupil was familiar with locus and graphs so he was asked in a similar way to deduce the equation of a circle. Only after strong hints did he notice the use of pythagoras even though this challenge followed the previous one.

In the development of the circle theorems, strong hints were needed in order that the pupil could make use of the external angle theorem for triangles (just proved) in the proof of the first theorem. And only in the final theorem on the angle in a semi-circle did he see without strong hints that it was only an application of the first theorem. He needed strong hints to use pythagoras in the solution of series 3 qu. 1, though he did not quite see how to use it.

Item 8 (M) Able to describe a strategy for solving simple equations of the sort $ax + b = c$ but cannot describe an overall strategy for every form of simple equation without using examples.

Required examples to describe strategies in finding the probability of an event.

Able to give a general method for drawing a graph.

Cannot follow the solution of $ax^2 + bx + c = 0$ with confidence.

In generalising $(a + b)^2$ to $(C + D + E)(E + C + D)$ (see appendix) needs to proceed in several steps 2,2 4,4,5,5, 8

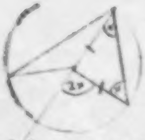
Cannot find a general rule for $(a5)^2$ (see appendix)

Able to work with algebra in straightforward cases like proof of $\sin^2 x + \cos^2 x = 1$.

Gave wrong answer to $(a + b)^2$, saying it was $a^2 + b^2$ at first, but was able to understand it correctly after careful explanation.

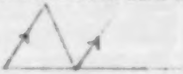
ITEM 9 (M) 10 (L)

Was able to miss out simple steps in mathematical arguments such as if $3x + 2 = x+1$ then $2x = 1$ and from the proof of the first circle theorem.

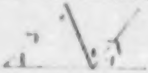


When the examiner had demonstrated that by splitting the diagram into two isoscles triangles the external angle theorem could be applied the pupil observed that it was 'the same for the other triangle' after his attention was drawn to it.

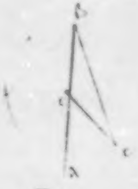
He could not make other, more telling logical jumps, however, even when given ample opportunity and in situations where pupils of high ability did make the jumps. For example: he needed every step of the proof of the external angle theorem, for a triangle where others commonly could see the proof after the parallel construction line had been drawn



or even when some angles were marked.



He could not miss out any steps in the proof of



$\hat{AOC} = 2\hat{OBC}$ from the first circle theorem, and the examiner had to restart when too much was assumed such as "$\triangle OBC$ is isosceles", or having justified $\triangle OBC$ as being isosceles, "Well that gives equal angles OBC and OCB, so remembering the result about triangles," (pointing to the result on the paper which had just been proved), "then $\hat{AOC} = 2\hat{OBC}$."

Similar opportunities were presented in several other aspects of the work.

ITEM 11 (M)

Whenever the pupil was unable to proceed with a mathematical process he tended to persist, but not try alternative methods from a position that had been validly reached. In the Series 3 problem 6, on missing numbers he restarted the problem each time a number failed. When the examiner demonstrated the results P1 to P4 and when a step was not understood, the examiner had to recommence the work from the beginning except in cases where it had not developed very far such as

$$(1) \quad 3x + 3y = 7$$

$$(2) \quad 2x - y = 6$$

Multiply (1) by 2 and (2) by 3

$$6x + 6y = 14$$

$$4x - 2y = 12$$

(Examiner) "Oh that's wrong, I should have multiplied by 3",

$$6x - 3y = 18$$

"That's right now . . . . .

ITEM 12 (M)

The pupil gave only one idea on each occasion that he was asked. Some of these havebeen listed in for ITEM 4. He was aware that alternatives were possible but could not discover them except when given strong hints, though he would not readily give up. In series 4 Qu 3, he would only think of straightforward additions. The examiner demonstrated three other methods (adding in pairs, reversing and adding and use of a formula). In series 4, Qu 4 the pupil could only perform $113^2 - 112^2$ by simple computation. The pupil enjoyed the $(113^2 - 112^2$ $(113 + 112)$solution, and appreciated that there were other less useful ways such as drawing squares and measuring; he suggested plotting graphs after strong hints, but could not find any other workable solutions himself here or elsewhere.

ITEM 13 (M)

His lack of ability to find alternative solutions himself made him unable to discover the most elegant, but he did show pleasure in the best solution presented such as in series 4 Qu.4 and the algebraic method for series, realising that the idea could be used for other addition (a further clue to ITEM 7). Pupils who scored lower than him on this challenge did not appreciate an elegant solution and would choose the one that they themselves had chosen or understood most easily.

ITEM 14 (M)

The pupil noticed simple slips of logic such as that reported on simultaneous equations for ITEM 11 and in cases where deliberate numerical slips were made. He corrected a deliberate slip in inequalities "$-3X \leq 5$ to $X \leq -{}^{5}/_{3}$" but only when the examiner asked him a second time if he was sure that the deduction was correct. He saw through the invalid logic concerning the 1,000,000 ancestors (see Appendix 1), but failed to detect incorrect logic in 62.5 x 62.5, saying that the result could be expected to be wrong since the methods were different. He was also prepared to accept that it was possible to prove $a \neq b \Rightarrow a = b$ after attempting to check through the algebra.

ITEM 15 (M)

To score L (Low) on this ITEM, a pupil would indicate indifference to overcoming difficulties and it was only due to the persistance of the examiner that the pupil was persuaded to gain proper understanding. To score H (High) a pupil would make determined efforts to concentrate and understand and give indications that he/she would work independently to gain complete understanding. This pupil showed an interest in the work that was done, and showed willingness to try again when something was not understood (this occurred in all of the proofs and methods demonstrated and reported earlier). In cases where constant failure occurred, such as quadratic equations there was an admission of defeat, however, and the pupil stated that it was not worthwhile to try again on his own after the testing session. He did, however, take away one of the problems (Series 3 Q5) because he would enjoy trying it again.

ITEM 16 (L)

The pupil took well over a minute to adjust from direct to reverse in the series 5 problems (Q. 1, 2, 3, 4) and admitted that he needed to spend time to get his thoughts straight.

ITEM 17 (M)

The pupil was able to remember, at the end of the session, the techniques which had been new and taught to him at the beginning. He demonstrated how to solve simultaneous equations but needed an example in order to demonstrate the algorithm. He remembered the proofs for the circle theorems and for the external angle theorem, but he had to give every step, rather than the essential details that other pupils who scored H (High) were inclined to do. He was able to adapt the method for simultaneous equations to a slight variation (1) $3X = 5-y$
(2) $X + 2y = 5$.

ITEM 19 (M)

He was not showing signs of great fatigue at the end but said that though he had enjoyed the session he was ready for a break. The examiner sensed most fatigue after the pupil had worked on unfamiliar topics and found that the pupil benefited from a short break between such topics or when a second attempt at an explanation was needed such as in the proof of the first circle theorem.

ITEM 20 (M)

The pupil showed interest in the series 3 puzzles and asked if he could make a copy of Q. 5 to try at home. He did not show the same willingness to spend extra time on all of the problems nor in resolving the difficulties that he had found in solving quadratic equations.

2. Pupil of known high ability HM5

This pupil was so quick with the work that virtually every challenge that the bank contained was covered.

ITEM 1 (H)

Answered correctly and quickly all of series 1a, 1b, 1c. Made an immediate start on all series 3 problems, and succeeded in solving them indicating the correct use of information. Succeeded on other examples too such as the fallacious proof of $45^\circ = 60^\circ$, having no trouble with a complex diagram.

ITEMS 2/3 (H)

Very quick to answer correctly all series 2 questions. No trouble with any series 3 questions where the essential information had to be extracted.

ITEM 4 (H)

Very quick to suggest ideas whenever asked. Strongest indication from the unfamiliar series 3 problems where he not only suggested a plausible method but actually solved the problems.

ITEM 5 (H)

Followed everything demonstrated with ease (even when the examiner deliberately made it difficult by deliberate mistakes necessitating fresh starts). The strong evidence came when the pupil could reproduce such things as the method for the general solution of quadratics, proof of the cosine rule for triangles, use of the method of induction, all of which had been new work and demonstrated in a way that was hard to follow.

ITEM 6 (H)

Many indications. In particular, solved all of the series 3 problems.

ITEM 7 (H)

In the unfamiliar proof of $\sin^2 x + \cos^2 x = 1$ spotted the use of pythagoras immediately as he did with the derivation of the equation of a circle. The Sine rule and Cosine rule were demonstrated early in the session and late in the session the examiner discussed another unfamiliar topic, that of vectors. The examiner drew a diagram and was about to ask the pupil if he could suggest how one might calculate the resultant of two vectors. The pupil, however, said 'this is an example where the sine and cosine rules will be applied', before the examiner had time to ask! The pupil also gave immediate answers when the examiner asked for examples where maths topics linked together such as a) Ratio and Fractions b) Linear programming and simultaneous equations, sets, inequalities, graphs, c) Inequalities and Sets.

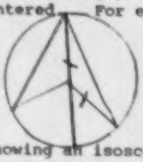
ITEM 8 (H)

The pupil gave an immediate generalisation of $(a+b)^2$ to (E+F+G) (E+G+H) so the examiner suspected prior knowledge and challenged him with the unfamiliar sin (A+B+C) after first teaching him the expansions of sin (A+B) and cos (A+B).

The generalisation was immediate. In finding a rule for $(X5)^2$ (stated in less specific terms of a digit followed by 5 squared like 25^2) he considered, first, $(X+\frac{1}{2})^2$ and found a rule by expanding the brackets and then found a general rule for $(X5)^2$ by multiplying the result by 100 i.e. "square X , add X , add .25 and then multiply by 100". He was able to follow and learn quickly all of the algebraic proofs, including the method of mathematical induction. When asked to describe a method of solution, such as the solution of quadratic equations he described it in general terms with no reference at all to particular examples.

ITEM 9/10 (H)

The ability of this pupil was detected early in the session so no attempt was made to plod step by step through any proofs. Many steps were assumed and no difficulties were encountered. For example in the first circle theorem proof the diagram was drawn



the equal radii were marked, showing an isosceles triangle. It was then remarked that the outside angle was equal to the two interior and opposite and thus twice as much as each. The same applied to the other triangle, hence the result. The pupil was asked if he understood and was asked to repeat the proof with full details, which he did. He was then able to prove the remaining circle theorems himself almost immediately, having made the jump of logic from the first theorem to its application with the others.

ITEM 11 (H)

Many examples of his flexibility were evident. One strong example came from the examiner's deliberate attempt to make the solution of the cosine rule hard to follow. It was commenced by drawing a diagram and pursuing the normal proof of the sine rule and even here deliberately making slips and restarting mid proof. The sine rule was proved and the examiner apologised, blaming a hazy memory and restarted from the original diagram taking a rather longwinded route and then having nearly finished the proof, said that he was not happy, looked it up in a text book and restarted again mid proof but with a 'neater' diagram labelled in a different way. The pupil was able to reproduce the best proof immediately.

ITEM 12 (H)

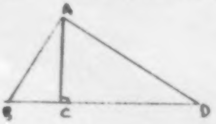
Able to find alternative ways to start series 3 problems when one method failed. For example in Q. 2 first tries to clear brackets and it leads him nowhere. He got annoyed, but restarted and saw that it was a quadratic in X. (The examiner had only introduced him to quadratics in this session - Item 8!) He then suggested that trial and error would work as well. In the series 4 problems gave at least two methods. When asked how to solve simultaneous equations he gave three methods - graph, algebraically (two ways), trial and error.

ITEM 13 (H)

In his solutions to series 3 he was clearly seeking a neat solution in each case. In discussion of the methods for quadratic equations (completing the square, formula, graph or factorisation) he discussed the merits of each method critically. He showed delight in the method of induction and discussed why he preferred it in the solution of 1+2+3+ --- +50 because even though it was longer than other methods it had application beyond one simple problem. He criticised it in that one needed to 'know the answer first'.

ITEM 14 (H)

Quick to spot the errors in all the illogical proofs, not only that they were there but also found what the errors were. Even the more complex problems on $a \neq b \Rightarrow a=b$ and $45^\circ = 60^\circ$ gave no difficulty, though in the latter example he spent time on the logic before spotting that the diagram was the cause of the fallacy. He was also quick to spot any deliberate errors that the examiner made during a demonstration such as, in the triangle ABC, sine B = $\frac{BC}{AB}$.



ITEM 15 (H)

This became clearest after the testing session. The examiner had thanked him for his help and hoped he would get in touch if he wanted help in maths at any time. The pupil had been in receipt of extra enrichment lessons and had become interested in Number Theory. He wrote to the examiner concerning Goldbach's conjecture, enquiring about any recent knowledge and offering his own proof. The communications are still continuing regarding this particular topic at the time of writing.

ITEM 16 (H)

Almost immediate responses to the Direct/Reverse series 5.

ITEM 17 (H)

All of the methods that were taught were remembered in curtailed structures. For example induction was remembered as 'Well we assume its true for some number. We add the next term to both sides and rearrange the formula so that it looks the same for one more. Then we see if it works for n = 1.' Proof of quadratic formula was remembered as 'Divide through by a. Take the constant term $\frac{c}{a}$ over. Complete the square. Find the square root and rearrange the answer.' All of the curtailed proofs could be followed by more detailed proofs.

ITEM 18 (H)

There was ample opportunity to make use of work that the pupil had previously learned. Techniques were always readily available. This is apparent, too, in the communication concerning Goldbach's Conjecture.

ITEM 19 (H)

This was a particularly long session (about four hours). The pupil was adamant that he did not want or need a break and showed no signs of fatigue whatsoever, indeed he wanted to carry on at the end.

ITEM 20 (H)

See Item 15. The pupil also learned some new work such as logs in order to make progress with his understanding of some aspects of number theory. His appetite seemed insatiable.

ITEM 22 (H)

The pupil admitted a strength in logical work and a weakness in geometric work. Compared with others this weakness was not evident, but in relative terms he spent a little longer on spatial work such as in finding the error in the $45^\circ = 60^\circ$ fallacy. He also said that he had taken up Latin at school because he enjoyed the way it was a logical subject.

Notes for Appendix 7

1) For example Thorndike, E.L., The Psychology of Arithmetic, New York, Macmillan 1922, Brownell, W.A., Pscychological Consideration in the Learning and Teaching of Arithmetic. The Teaching of Arithmetic, the tenth yearbook of The National Council of Teachers of Mathematics, New York, Teachers College, Columbia University 1935. Newell, A., Simon, H.A., Human Problem Solving, Englewood Cliff, N.J., Prentice Hall, 1972.

2) It was possible to use O-level scores rather than grades because of the co-operation of the Oxford Delegacy of Local Examinations.

3) Based on 1870 cases representing a sample from the whole ability range in mathematics.

4) This means that each of these variables was in the regression formula, giving its own independent contribution to the predicted score.

5) We can estimate the percentage variance explained by a variable by squaring the correlation coefficient. Thus SR explains about 11% of the variance in O-level mathematics score when its relationship is considered in isolation from other variables.

6) For example organisers connected with Schools Council Programmes throughout the country, the Nottinghamshire project for the provison for gifted children, the Royal Society Masterclasses.

7) To preserve the confidentiality of the identity of these pupils no names of organisations has been given in the table and the identity codes used later will not be related to the position of the pupil in this list.

8) For explanation of this procedure see Appendix 1.

9) The Seventh Mental Measurements Yearbook, Ed. Oscar K. Buros, Gryphon Press 1972.

10) Catalogue of Tests for Educational Guidance and Assessment, NFER Publishing Company (published annually).

11) Unlike the other items, these were the expected levels, rather than defined retrospectively. As will be discussed it was not possible to obtain convinving conclusions on this item.

APPENDIX 8

Teachers constructs in mathematics

As can be seen in Chapter Twelve, the main focus for work in Mathematics was on detailed classroom observation. The full construct study described in Chapter Eleven was not therefore attempted and a simple study which did not refer to the critical triad approach was substituded. Constructs were elicited from six teachers. Pupils ratings on these constructs were then correlated with their background test score. A high correlation would identify a construct to which, if a high degree of agreement between test and teacher judgement is sought, the teacher should attend in forming his/her overall judgement about pupils. The pupils ratings on the same constructs were also correlated with the teacher's overall judgement. High correlations indicated constructs to which the teacher did attend.

Results for each teacher are presented separately below. A pupil who fits the description in the tables would score 5 on the teachers ratings. A pupil at the opposite pole would score 1.

TABLE 1

MATHEMATICS CONSTRUCTS FOR TEACHER T041

	correlated with:	
Description of construct	Test Score	Teacher Selection
Good presentation of work	-.05	.31
High National Ability	.30	.67
High Accuracy of work	.17	.61
Neat presentation	.02	-.04
Regular attendance at school	.09	.24
Achieving potential	.44	.75
Strong personality	.24	.15
Good overall effort	-.10	.21
Clear thinking	.39	.56
Good understanding of work	.35	.63
Overconfident	.13	.01
Good organisation of work	.32	.44
Hard working	.13	.23
Talkative in class	-.16	-.28

TABLE 2

MATHEMATICS CONSTRUCTS FOR TEACHER T042

Description of construct	correlated with: Test Score	correlated with: Teacher Selection
High Natural ability	.61	.51
Hard working	.24	.31
Good presentation	.31	.33
Attentive in class	.25	.32
Takes work seriously	.31	.35
Achieves to maximum of his/her potential	.14	.11
Highly motivated	.16	.14
Good long term memory	.67	.62
Easily distracted	.29	.49
Asks questions in class	.31	.52
Does more than the work set	.29	.47
Good exam performance	.60	.62
Hands in homework on time	-.05	-.01
Coasts along	.09	.12
Highly dedicated	.52	.39

TABLE 3

MATHEMATICS CONSTRUCTS FOR TEACHER T05

Description of construct	correlated with: Test Score	correlated with: Teacher Selection
Good general maths ability	.32	.50
Good oral ability	.21	.18
Good written work	.05	.41
Confident	.30	.30
Careless	.36	-.15
Good Memory	.38	.40
Conscientious	-.05	.55
Achieves potential	-.21	-.18
Quiet in class	-	-
Tidy work	-.03	.02

TABLE 4

MATHEMATICS CONSTRUCTS FOR TEACHER T09

Description of construct	correlated with: Test Score	Teacher Selection
Interested in maths	.40	.72
Keen on maths	.36	.64
Mature	.11	.56
Tries Hard	.13	.58
Serious Attitude	.19	.38
Gets very confused	-.20	-.80
Disruptive	-.06	-.55
Keen to please	.13	.42
Untidy work	.21	.09
Casual Attitude	-.03	-.40
Co-operative	.05	.43
Strongly Motivated	.30	.45
Fast worker	.44	.51
Accurate worker	.18	.66
Overconfident	.30	.28

TABLE 5

MATHEMATICS CONSTRUCTS FOR TEACHER T12

Description of construct	correlated with: Test Score	Teacher Selection
Realises true potential	.29	.04
Industrious	.15	.47
Good at geometry	.21	.60
Good overall natural maths ability	.48	.83
Under confident	-.06	-.23
Good orally	.28	.56
Disruptive	-.13	.16
Able to make logical jumps	.55	.69
Good mental arithmetic	.56	.65
Good at following explanations	.54	.55
Increasing rate of progress	-.20	.28

Table 6 gives the constructs for the teacher in School 00. No correlations are reported for this teacher who based her nominations on checklist useage, so that the correlation approach could be misleading.

TABLE 6

MATHEMATICS CONSTRUCTS FOR TEACHER TOO

Description of construct	correlated with: Test Score	Teacher Selection
Good concentration		
High overall achievement		
High natural ability		
Good effort		
Achieving potential		
Strong interest		
Tidy work		
Makes Logical jumps		
Thinks logically		
Contributes in class		
Quiet		
Sensible		
Asks Questions		
Picks up a concept quickly		
Listens in class		
Gets on with work		
Keen		
Reliable		
Accurate		
Shows pleasure in understanding		

Multiple regression analysis was carried out with the teacher selection variable (1 for pupils selected, 0 for pupils not) as the dependent variable. This was done in order to allow for the intercorrelations between the construct ratings and isolate the constructs that were most strongly linked to teacher selection. Table 7 summarises the findings.

TABLE 7

VARIABLES THAT MADE SIGNIFICANT CONTRIBUTIONS TO THE VARIANCE WHEN THE TEACHER SELECTION VARIABLE IS SUBJECTED TO MULTIPLE REGRESSION ANALYSIS

TEACHER	% VARIANCE EXPLAINED BY REGRESSION FORMULA	VARIABLES CONTRIBUTING SIGNIFICANTLY TO THE VARIANCE	CORRELATION WITH TEACHER SELECTION	CORRELATION WITH TEST SCORE
T041	54%	Achieving potential	0.75	0.44
T042	43%	Good long term memory	0.62	0.67
		Asks questions in class	0.52	0.31
T05	43%	Conscientious	0.55	-0.05
T09	63%	Gets very confused	-0.80	-0.20
T12	78%	Natural maths ability	0.83	0.48
		Disruptive	0.16	-0.13
		Underconfident	-0.23	-0.06
		Rate of progress	0.28	-0.20

We can see that in all cases the amount of variance explained is quite high. Tables 1 to 6 illustrate, as we discovered from our informal interviews, that teachers tended to talk in rather general terms when describing the mathematical ability of their pupils. The fact that only a few of the constructs (often only one) contributed to the prediction equation tends to confirm the fact that a teacher could not take a large number of factors into account when making decisions about pupils' abilities. It may be that some of the constructs, such as 'natural maths ability' was associated with a general feeling concerning the child's ability on the whole range of mathematical aptitudes.

In some cases, such as T042 there is a high correlation with a particular variable and both teacher selection and test score which indicates that whatever the teacher meant, by, in this case 'long term memory' was used validly in the identification process.

There is evidence however that some teachers may use inappropriate criteria as single disqualifying one (for example T05) where we see that the dominant construct in the identification process had very little to do with test score. This is of considerable concern because of the intercorrelation of ratings on the inappropriate variable and the teachers rating on what seems on face value to be more appropriate.

For example correlation on ratings from Teacher T05 are: 'Conscientious' with 'Good general ability' 0.50, 'Conscientious' with 'Good memory' 0.45. This seems to imply that conscientiousness to the teacher, is strongly associated with ability.

From this small scale study we have further evidence to add to what we suspected from our work in the first year of the project. As a result of insufficient detailed knowledge or memory of their pupils' abilities in the subject teachers may tend to have general impressions regarding abilities and attitudes. Some teachers can put distracting constructs to one side and make judgements on generalities concerning 'ability' or 'potential' which are relevant but which still must lead to very coarse strategies. Others are distracted to some extent by attitude and motivation to a level that influences judgements of ability.

TABLE 8

CATEGORISATION OF CONSTRUCTS ASSOCIATED WITH EACH TEACHER

TEACHER	GENERAL MATHS ABILITY	SPECIFIC MATHS ABILITY	ATTITUDE/ MOTIVATION/ PERSONALITY	PRESENTATION OF WORK	OTHER
T041	5	0	4	3	2
T042	3	1	9	2	0
T05	4	1	3	2	0
T09	3	0	11	1	0
T12	4	4	3	0	0
T00	6	3	10	1	0
TOTALS	25	9	40	9	2

Table 8 shows that the majority of ways in which pupils are described by their teachers fall into the general ability or attitude/motivation/personality categories. Teacher T042 and T09 notice many more characteristics in the latter category. Teacher T00 is of particular interest. She had been concerned with identification by checklist and had carried out successful work in the second stage of that activity. The records were kept on paper and we see that her memory of detailed strengths and weaknesses of her pupils were not present when the construct ellicitation work was done. No doubt her memory of pupils who scored high on many checklist items was accurate, but one must conclude that without the accurate records she would not have been able to make clear judgements about many pupils in the middle band of her class with some strengths and some weaknesses, neither would she be able to cater for their individual needs in more than a general way.

BIBLIOGRAPHY

A.P.U., 1978, Assessment of Performance Unit, Science Progress Report, 1977-78. London, H.M.S.O.

A.P.U., 1978, Language Performance, London, H.M.S.O.

Argyle, J.M. & McHenry, R., 1971, Do Spectacles really affect our judgements of intelligence?, Brit. J. of Soc. and Clin. Psychol, 10.

Astin, H.S., 1975 , Sex differences in mathematical and scientific Precocity, J. Special Educ., 9, 1.

Barbe, W.B. & Renzulli, J.S., 1975, Psychology and Education of the Gifted, New York, Irvington.

Barnes, D., 1980, Situated Speech Strategies. Aspects of the Monitoring of Oracy, Educ. Rev. 32, No. 2.

Bender, M.P., 1974, Provided versus elicited constructs, Brit. J. Soc. Clin. Psychol. 13.

Bennett, G.K., Seashore, H.G. & Wesman, A.G., 1974, 5th Manual for the Differential Aptitude Tests, New York, Psychological Corporation.

Bridges, S.A., 1969, Gifted Children and the Brentwood Experiment, London, Pitman.

Bridges, S.A., 1975, Gifted Children and the Millfield Experiment, London, Pitman.

Brookover, W.B., Erikson, E.K. & Joiner, L.M., 1967, Self-Concept of Ability and School Achievement III, East Lansing, Michigan Stage Univ.

Brownell, W.A., 1935, Psychological Consideration in the Learning and Teaching of Arithmetic in The Teaching of Arithmetic - the 10th yearbook of the National Council of Teachers of Mathematics, New York, Teachers College, Columbia University.

Buros, O.K., 1972, The Seventh Mental Measurement Yearbook, New York, Gryphon Press.

Burt, C., 1962, General Introduction: The Gifted Child, The Yearbook of Education (Ed. Bereday and Lauerys), Evans.

Butcher, H.J., 1966, Sampling in Educational Research, Manchester, Manchester University Press.

Cattell, R.B., 1937, Measurement versus Intuition in Applied Psychology, Character & Personality 6.

Chronback, L.E. & Meehl, P.E., 1955, Construct Validity in Psychological Tests, Psychological Bulletin, 52.

City of Birmingham Education Committee, 1979, Gifted and Outstanding Children, Birmingham.

Clarke, G., 1981, Guidelines for the recognition of more able pupils (Draft), London, Schools Council.

Clarke, G. & Cutland, P., 1977, Identifying for Provision, Cleveland County Council.

Cohen, L. & Manion, L., 1981, Perspectives on Classrooms and Schools, London, Holt Education.

Cook, M., 1979, Perceiving others, London, Methuen.

Cornish, R.L., 1968, Parents' Teachers' & Pupils' perception of the gifted child's ability, Gifted Child Quarterly, 12.

Cunningham, C., Thomson, B., Alston, H. & Wakefield, J., Winter 1978, Use of S.O.I. Abilities for Prediction, The Gifted Child Quarterly.

Devon C.C., 1977, Find the Gifted Child. Devon Education Dept.

Dudeney, H.E., 1967, Puzzles and Curious Problems, London, Fontana.

Dunn, A., 1980, Mathematical Bafflers, New York, Dover.

Earnshaw, H.G., 1974, Criterion and Assessment in English, Associated Examining Board.

Easterby-Smith, M., 1981, in Shaw M.L.G. (ed.) Recent advances in Personal Construct Technology, London, Academic Press.

Eggleston, J.F., Galton, M.L., & Jones, M.E., 1976, Processes and Products of Science Teaching, London, Macmillan.

Elashoff, J.D. & Snow, R.E., (Eds.), 1971, Pygmalion Reconsidered, Worthington, Ohio, Charles A. Jones.

Entwistle, N., 1968, Academic Motivation and School Attainment, Brit. J. of Educ. Psychol, 38, 2.

Entwistle, N. & Wilson, J., 1977, Degrees of excellence, London, Hodder & Stoughton.

Eysenck, H.J., 1967, Intelligence Assessment, Brit. J. Educ. Psychol. 37.

Fransella, F. & Bannister, D., 1977, A manual for Repertory Grid Techniques, London, Academic Press.

Freeman, J., 1981, Some Emotional aspects of Giftedness. In The Gifted Child at School, (Ed. Grubb, D.W.H.), Oxford Society for Applied Studies in Education.

Frohlichstein, J., 1967, Mathematical Fun, Games and Puzzles, New York, Dover.

Gallagher, J.J., 1966, Research Summary on Gifted Child Education, Illinois, Dept. of Program Development for Gifted Children.

Getzels, J.W. & Jackson, P.W., 1958, The Meaning of Giftedness, Phi. Delta Kappan, 40, 2.

Getzels, J.W. & Jackson, P.W., 1962, Creativity and Intelligence, New York, Wiley.

Greenberg, L., 1955, Critique of classic methods of identifying gifted children, School Rev. 63.

Greer, A., 1973, A Complete O level Mathematics, London, S. Thornes.

Grubb, D.W.H., (Ed.), 1981, The Gifted Child at School, Oxford Society for Applied Studies in Education.

Guilford, J.P., 1957, Three Facets of the Intellect, American Psychologist Vol.14, No. 8, Aug. 1959, reprinted in Barbe, W.B. and Renzulli, J.S., Psychology and Education of the Gifted, Irvington.

Guilford, J.P., Howpfrer, K. & Peterson, H., 1965, Predicting Achievement in 9th Grade Mathematics from Measures of Intellectual Aptitude Factors, Ed. and Psych. Measurement, Vol.XXV, No. 3.

Guilford, J.P. & Fruchter, B., 1978, Fundamental Statistics in Psychology and Education, (6th Ed.), New York, McGraw Hill.

Guilford, J.P., 1967, The Nature of Human Intelligence, New York, McGraw Hill.

H.M.I., 1979, Aspects of Secondary Education, London, H.M.S.O.

H.M.I., 1977, Gifted Children in Middle and Comprehensive Secondary Schools, London, H.M.S.O.

H.M.I., 1964, The Eighth Report of the Secondary Schools Examination Council on the Examining of English Language, H.M.S.O.

Hall, E., 1978, Using Personal Constructs, Nottingham, Univ. of Nottingham, School of Education.

Harben, W., Darwin Sr., A., & Murphy, M., 1977, Match and Mismatch - Raising Questions, London, Oliver and Boyd.

Hargreaves, D., 1972, Interpersonal Relations and Education, London, Routledge & Kegan Paul.

Holbrook, D., 1961, English for Maturity, Cambridge, Cambridge University Press.

Hollingworth, L.S., 1942, Children Above 180 I.Q., New York, Yonkers-on-Hudson.

Hopkins, K.D. & Glass G.V., 1978, Basic Statistics for the Behavioural Sciences, New York, Prentice Hall.

Hoyle, E. & Wilks, J., 1980, Gifted Children and their Education, London, H.M.S.O.

Jacobs, J.C., 1971, Effectiveness of teacher and parent identification of gifted children as a function of school level, Psychol. in the Schools, 8.

Jenkins, E. & Whitfield, R.C., 1974, Readings in Science Education, London, McGraw Hill.

Kelly, G., 1955, The Psychology of Personal Constructs, New York, Norton.

Kerry, T., 1980, Teaching Bright Children, Teacher Education Project, Nottingham School of Education.

Kerry, T., 1981, Teaching Bright Pupils in Mixed Ability Classes. D.E.S. Teacher Education Project, London, Macmillan.

Krutetskii, V.A., (Tr. J. Teller), 1976, The Psychology of Mathematical Abilities in Schoolchildren, Chicago, University of Chicago Press.

Kulm, G. & Bussman, H., 1980, A Phase-Ability Method of Mathematical Problem Solving, Journal for Research In Mathematical Education, Vol.II, No. 3.

Kulm, G., April 14th, 1980, Identification and Analysis of Mathematical Problem Solving Processes, Vancouver BC, Paper presented to Research Council for Diagnostic and Prescriptive Mathematics.

L.A.T.E., 1965, Assessing Compositions, London, London Association for the teaching of English.

Laycock, S.R., 1957, Gifted Children: A Handbook for the Classroom Teacher, Toronto, Copp-Clark.

Levinson, B.M., 1956, Rethinking the selection of the intellectually gifted, Psychol. Reports, 2.

Locksby, A., Bargida, E., Brekke, N. & Hepburn, C., 1980, Sex Stereotypes and Social Judgement, J. of Personality and Soc. Psychol, 39.

Lovell, K. & Shields, J.B., 1967, Some Aspects of the Study of the Gifted, Brit. J. Ed. Psych. 37.

Marland, S.P. Jr., 1972, Education of the Gifted and Talented, Washington D.C., Report to Congress by the U.S. Commissioner for Education, Vol.1.

Martinson, R.A., 1975, The Identification of the Gifted and Talented, Virginia, Council for Exceptional Children.

Massey, A.J., 1978, The Relationship between O- and A-level Nuffield Physics Results, Educational Research, Vol.20, No. 2.

Meehl, P.E., 1961, Logic for the Clinician, Contemp. Psychol. 7.

Meehl, P.E., 1954, Clinical Versus Statistical Prediction, Minneapolis, Univ. of Minnesota Press.

Minium, E.W., 1970, Statistical Reasoning in Psychology and Education, New York, Wiley.

Modern Mathematics for Schools, 1972, Scottish Mathematics Group, London, Blacki Chambers.

Murphy, R.J.L., 1981, O level grades & teachers' estimates as predictors of the A level results of UCCA applicants, Brit. J. Educ. Psychol. 51.

N.F.E.R. Publishing Company, Catalogue of Tests for Educational Guidance and Assessment.

Newbold, D., 1977, Ability Grouping, The Banbury Enquiry, Slough, N.F.E.R. Publishing Co.

Newell, A. & Simon H.A., 1972, Human Problem Solving, Englewood Cliff, N.J., Prentice Hall.

Nie, H., Hull, C., Jenkins, J., Steinbrenner, K. & Bent, D., 1975, Statistical Package for the Social Sciences, New York, McGraw Hill.

Nisbet, J.F., 1895, The Insanity of Genius, London, Deloa Nore Press.

Northrop, P.N., 1944, Riddles in Mathematics, London, Pelican.

Ogilvie, E., 1980, The Schools Council Curriculum Enrichment Project, In The Gifted Child, (Ed. Povey), London, Harper and Row.

Ogilvie, E., 1973, Gifted Children in Primary Schools, London, Macmillan.

Oskarrson, 1975, The Relationship Between Foreign Language Proficiency and various Psychological Variables, Stutgart, 4th International Congress of Applied Linguistics.

Parkyn, G.W., 1948, Children of High Intelligence, a New Zealand Study, New Zealand Council for Educational Research, Oxford University Press.

Pegnato, C.W. & Birch, J.W., 1959, Locating Gifted Children in High Schools: A comparison of methods, Exceptional Children, Vol.25. No. 7.

Pimsleur, 1964, Underachievement in Foreign Language Learning, International Review of Applied Linguistics, Vol.2.

Pope, M.L., & Keen, T.R., 1981, Personal Construct Psychology & Education, London, Academic Press.

Postlethwaite, K.C. & Denton, F.C., 1978 Streams for the Future?, Banbury, Pubansco.

Povey, R., (Ed.), 1980, Educating the Gifted Child, London, Harper & Row.

Purusrothoman, M., 1976, Secondary Mathematics Item Bank, Slough, N.F.E.R. Publishing Co.

Renzulli, J.S., Reis, S.M. & Smith, L.H., May, 1981, The Revolving Door Model, Phi. Delta Kappan.

Renzulli, J. & Smith, L.H., 1977, Two approaches to the identification of gifted students, Exceptional Children.

Renzulli, J.S. & Hartman, R.K., November, 1971, Scale for Behavioural Characteristics of Superior Students, Exceptional Children.

Renzulli, J.S., Hartman, R.K. & Callahan, C.M., November, 1971, Teacher Identification of Superior Students, Exceptional Children.

Rogers, C., 1982, A Social Psychology of Schooling, London, Routledge & Kegan Paul.

Rosenthal, R. & Jacobson, L., 1968, Pygmalion in the Classroom, New York, Holt, Reinhard and Winston.

Rutter, M., Maugham, B., Mortimore, P. & Onslow, J., 1979, Fifteen Thousand Hours, London, Open Books.

Sarbin, T.R., Taft, R. & Bailey, D.E., 1960, Clinical Inference & Cognitive Theory, New York, Holt, Rinehart & Winston.

Sarbin, T.R., 1944, The Logic of Prediction in Psychology, Psychol. Rev. 51.

Schofield, B., Murphy, P., Johnson, S. & Black, P., Science in Schools Age 13: Report No 1., London, (A.P.U.), H.M.S.O.

School Mathematics Project, Westfield College, Cambridge University Press.

Semper, E. & Dutton, P., 1980 The Next Frontier, Paper 20, Sheffield Regional Centre for Science & Technology.

Shertzer, B., 1960, Working with superior children, Chicago, S.R.A.

Shields, J.B., 1973, *The Gifted Child*, Slough, N.F.E.R. Publishing Company.

Siegel, S., 1956, *Non Parametric Statistics*, New York, McGraw Hill.

Solomon, R., 1979, *Identifying Gifted Children*, J. of Appl. Educ. Stud., 8, 2.

Songuist, J.A. & Morgan, J.N., 1964, *The Detection of Interaction Effects*, Mono No. 35 Survey Research Centre. University of Michigan.

Stevens, F., 1970, *English and Examinations*, London, Hutchinson Educational.

Straker, A., 1981, *Mathematics for Able Pupils*, London, Schools Council.

Sutton, C.R. & Hayson, J.T., 1974, *The Art of the Science Teacher*, London, McGraw Hill.

Tempest, N.R., 1974, *Teaching Clever Children 7-11* London, Routledge & Kegan Paul.

Terman, L.M., (et al), *Genetic Studies of Genius*, (Vol.1, Terman, 1925. Vol.2, Cox, 1926. Vol 3, Burks, Jensen and Terman, 1930. Vol.4, Terman and Oden, 1947. Vol.5, Terman and Oden, 1959.), Stanford University Press.

Thompson, J.E., 1961, *Puzzle Playtime*, London, Wheaton & Co.

Thorndike, E.L., 1922, *The Psychology of Arithmetic*, New York, Macmillan.

Thurstone, L.L., 1952, *The Criterion Problem in Personality Research, Psychometric Lab., Rep. No. 78*, University of Chicago.

Thurstone, T.G., 1941, *Primary Mental Abilities of Children*, Educational and Psychological Measures, Vol.1, No. 12.

Torrance, E.P., 1965, *Gifted Child in the Classroom*, New York, Macmillan.

Torrance, E.P., 1961, *Problems of Highly Creative Children*, Gifted Child Quarterly, 5.

Vernon, P.E., 1969, *Intelligence and Cultural Environment*, London, Methuen.

Vernon, P.E., 1961, *The Structure of Human Abilities*, London, Methuen.

Vernon, P.E., 1957, Secondary School Selection, London, Methuen.

Wilkinson, A. & Hanna, P., 1980, The Development of Style in Children's Writing, Educational Review, 32, No. 2.

Wilkinson, A. & Hanna, P., 1979, Style in Children's Writing, Crediton Project in Language of Learning, Vol.1, No. 2.

Wilkinson, A., 1979, Assessing Language Development: The Crediton Project, Language for Learning, Vol.1, No. 2.

Willmott & Nuttall, 1975, The Reliability of Examinations at 16+, London, Schools Council Research Studies.

Wilson, C., 1963, Using test results and teacher evaluation in identifying gifted pupils, Personnel & Guidance, 41.

Witty, P., (Ed.), 1951, The Gifted Child. The American Association for Gifted Children, Boston, Heath & Co.

Wood, R. & Napthali, N.A., 1975, Assessment in the Classroom: What do teachers look for?, Educ. St., 1,3.

Yorke, D.M., 1978, Repertory Grids in Educational Research, Brit. Educ, Res. J., 4, 2.

Youngman, M.B., 1979, Some determinants of early secondary school performance, B.E.R.A. Conference Paper, Sept. 1979.